

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:39:51 ; Search time 80.5 Seconds  
(without alignments)  
108.141 Million cell updates/sec

Title: US-10-705-819B-4

Perfect score: 89

Sequence: 1 PRSDRPVSYTMRTRS 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03:\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	662	1 VGLF CDVO	P12569 canine dist
2	89	100.0	662	2 O89327	O89327 canine dist
3	89	100.0	662	2 Q9DX22	Q9DX22 canine dist
4	79	88.8	266	2 O55336	O55336 canine dist
5	78	87.6	266	2 O55333	O55333 canine dist
6	78	87.6	266	2 O55335	O55335 canine dist
7	78	87.6	266	2 O55338	O55338 canine dist
8	78	87.6	662	2 Q91KN3	Q91KN3 canine dist
9	78	87.6	662	2 Q6TV24	Q6TV24 canine dist
10	78	87.6	662	2 Q6TV32	Q6TV32 canine dist
11	71	79.8	266	2 O55337	O55337 canine dist
12	70	78.7	662	2 Q6SVY0	Q6SVY0 canine dist
13	64	71.9	662	2 Q68PH9	Q68PH9 canine dist
14	64	71.9	662	2 Q6T1B5	Q6T1B5 canine dist
15	64	71.9	662	2 Q9YKL7	Q9YKL7 canine dist
16	56	62.9	266	2 O55334	O55334 canine dist
17	47	52.8	157	2 Q657A9	Q657A9 oryza sativ
18	45	51.7	297	2 Q82G57	Q82G57 streptomyc
19	45	50.6	331	2 Q9VR31	Q9VR31 drosophila
20	44	49.4	167	2 Q6ZBR3	Q6ZBR3 oryza sativ
21	44	49.4	465	2 Q6TVH9	Q6TVH9 bovine papu
22	44	49.4	485	2 Q8PNQ3	Q8PNQ3 xanthomonas
23	43	48.3	111	2 O8SG62	O8SG62 oryza sativ
24	43	48.3	201	2 Q21133	Q21133 caenorhabdi
25	43	48.3	496	2 Q6MQ19	Q6MQ19 bdellovibri
26	43	48.3	530	2 Q759W6	Q759W6 ashbya goss
27	43	48.3	559	2 Q6CD99	Q6CD99 yarrowia li
28	43	48.3	658	2 Q76LU8	Q76LU8 oryzae lat
29	43	48.3	984	2 Q751J9	Q751J9 ashbya goss
30	42	47.2	104	2 Q9URH2	Q9URH2 saccharomyc
31	42	47.2	166	2 Q84YV9	Q84YV9 oryza sativ

## RESULT 1

VGLF CDVO STANDARD; PRT; 662 AA.

AC P12569; Q65991; AC  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
GN Fusion glycoprotein F1].  
GN Name=F;  
OS Canine distemper virus (strain Onderstepoort) (CDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11233;  
RN [1]  
RX MEDLINE=88129050; PubMed=3433924; DOI=10.1016/0168-1702(87)90009-8;  
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;  
RT "The nucleotide sequence of the gene encoding the F protein of canine  
distemper virus: a comparison of the deduced amino acid sequence with  
other paramyxoviruses.";  
RL Virus Res. 8:373-386(1987).  
RN [2]  
RX MEDLINE=93227636; PubMed=8470428; DOI=10.1016/0264-410X(93)90285-6;  
RA Wild T.F., Bernhard A., Spehner D., Villevial D., Drillean R.;  
RT "Vaccination of mice against canine distemper virus-induced  
encephalitis with vaccinia virus recombinants encoding measles or  
canine distemper virus antigens.";  
RL Vaccine 11:438-444(1993).  
CC -!- FUNCTION: This protein directs fusion of viral and cellular  
membranes.  
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked.  
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M21849; AAA42878.1; -;  
EMBL; X65509; CAA46481.1; -;  
PIR; JS0321; VGNZCD.

PIR; S21382; S21382.  
HSSP; P04849; 1SVF.

DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; Fusion\_gly; 1.

KW Envelope protein; Fusion protein; Glycoprotein; Signal; Transmembrane.  
FT SIGNAL 1 ? 662 Fusion glycoprotein F0.

CHAIN 1 ? 662 Fusion glycoprotein F0.

## ALIGNMENTS

32 42 47.2 265 2 Q62J53  
33 42 47.2 265 2 Q63SU3  
34 42 47.2 403 2 Q9PSS5  
35 42 47.2 511 2 Q6A573  
36 42 47.2 614 2 Q52516  
37 42 47.2 638 2 Q8RL93  
38 42 47.2 820 2 Q60585  
39 42 47.2 1302 1 DOC7 HUMAN  
40 42 47.2 1520 2 Q6ZV32  
41 41 46.1 114 2 Q8L3S7  
42 41 46.1 114 2 Q9SLC8  
43 41 46.1 133 2 Q9PID2  
44 41 46.1 207 2 Q88S93  
45 41 46.1 293 2 Q6X3U0

```
FT CHAIN ? 224 Fusion glycoprotein F2.
FT CHAIN 225 662 Fusion glycoprotein F1.
FT TRANSMEM 606 629 Potential.
FT DISULFID 180 307 Linkage between F2 and F1 (Potential).
FT CARBOHYD 62 62 N-linked (GLNAC. .) (Potential).
FT CARBOHYD 141 173 N-linked (GLNAC. .) (Potential).
FT CARBOHYD 173 173 N-linked (GLNAC. .) (Potential).
FT CARBOHYD 179 179 N-linked (GLNAC. .) (Potential).
FT CARBOHYD 179 179 R -> K (in Ref. 2).
FT CONFLICT 3 3 D -> N (in Ref. 2).
FT CONFLICT 140 140 N -> S (in Ref. 2).
FT CONFLICT 152 152 I -> M (in Ref. 2).
FT CONFLICT 171 171 A -> V (in Ref. 2).
FT CONFLICT 174 174 L -> H (in Ref. 2).
FT CONFLICT 662 662
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797803F0 CRC64;

Query Match 100.0%; Score 89; DB 1; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
Db |||||
50 PRSDRPVSYTMNRTS 66

RESULT 2
ID 089327 PRELIMINARY; PRT; 662 AA.
AC 089327;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP MEDLINE=93174978; PubMed=8438593;
RX Sidhu M.S., Husar W., Cook S.D., Dowling P.C., Udem S.A.;
RT "Canine distemper terminal and intergenic non-protein coding
RT nucleotide sequences: completion of the entire CDV genome sequence.";
RL Virology 193:66-72(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Sidhu M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF014953; AAC26994.1; -.
DR HSP; P04849; ISVF.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72951 MW; 80B144C6B9801898 CRC64;

Query Match 100.0%; Score 89; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
Db |||||
50 PRSDRPVSYTMNRTS 66

RESULT 3
Q9DXZ2
```

```
ID Q9DXZ2 PRELIMINARY; PRT; 662 AA.
AC Q9DXZ2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20499096; PubMed=11044118;
RX DOI=10.1128/JVI.74.22.10737-10744.2000;
RA Gassen U., Collins F.M., Duprex W.P., Rima B.K.;
RT "Establishment of a rescue system for canine distemper virus.";
RL J. Virol. 74:10737-10744(2000).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
CC EMBL; AF305419; AAG30919.1; -.
DR HSP; P04849; ISVF.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72988 MW; 9C5C1398CAE7B4C CRC64;

Query Match 100.0%; Score 89; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
Db |||||
50 PRSDRPVSYTMNRTS 66

RESULT 4
ID 055336 PRELIMINARY; PRT; 266 AA.
AC 055336;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF026242; AAB88267.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
FT NON TER 266 266
SQ SEQUENCE 266 AA; 29135 MW; 16A6FEA46AEC1DAC CRC64;

Query Match 88.8%; Score 79; DB 2; Length 266;
Best Local Similarity 88.2%; Pred. No. 2.4e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

Qy 1 PRTSDRPVSYTMNRTS 17
Db 50 PRTSDRPVHVTMNRTR 66

RESULT 5
O55333 PRELIMINARY; PRT; 266 AA.
ID O55333;
AC O55333;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF026239; AAB88264.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON_TER 266
SQ SEQUENCE 266 AA; 29164 MW; C8C813FC0A366375 CRC64;

Query Match 87.6%; Score 78; DB 2; Length 266;
Best Local Similarity 88.2%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17
Db 50 PRTSDRPVHVTMNRTR 66

RESULT 6
O55335 PRELIMINARY; PRT; 266 AA.
ID O55335;
AC O55335;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF026241; AAB88266.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON_TER 266
SQ SEQUENCE 266 AA; 29164 MW; C8C813FC0A366375 CRC64;

Query Match 87.6%; Score 78; DB 2; Length 266;
Best Local Similarity 88.2%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17
Db 50 PRTSDRPVHVTMNRTR 66

RESULT 7
O55338 PRELIMINARY; PRT; 266 AA.
ID O55338;
AC O55338;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF026244; AAB88269.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON_TER 266
SQ SEQUENCE 266 AA; 29389 MW; BA762A02F1F39DEC CRC64;

Query Match 87.6%; Score 78; DB 2; Length 266;
Best Local Similarity 88.2%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17
Db 50 PRTSDRPVHVTMNRTR 66

RESULT 8
O91KN3 PRELIMINARY; PRT; 662 AA.
ID O91KN3;
AC O91KN3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Onderstepoort;
EX MEDLINE=21306344; PubMed=11413309;
RA DOI=10.1128/JVI.75.14.6418-6427.2001;
RA von Messling V., Zimmer G., Haas L., Cattaneo R.;
RT "The hemagglutinin of canine distemper virus determines tropism and
cytopathogenicity.";
RL J. Virol. 75:6418-6427(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

```

SQ SEQUENCE 266 AA; 29191 MW; 7A80B29F09E872EE CRC64;

Query Match 87.6%; Score 78; DB 2; Length 266;
Best Local Similarity 88.2%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17
Db 50 PRTSDRPVHVTMNRTR 66

RESULT 7
O55338 PRELIMINARY; PRT; 266 AA.
ID O55338;
AC O55338;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF026244; AAB88269.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON_TER 266
SQ SEQUENCE 266 AA; 29389 MW; BA762A02F1F39DEC CRC64;

Query Match 87.6%; Score 78; DB 2; Length 266;
Best Local Similarity 88.2%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17
Db 50 PRTSDRPVHVTMNRTR 66

RESULT 8
O91KN3 PRELIMINARY; PRT; 662 AA.
ID O91KN3;
AC O91KN3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Onderstepoort;
EX MEDLINE=21306344; PubMed=11413309;
RA DOI=10.1128/JVI.75.14.6418-6427.2001;
RA von Messling V., Zimmer G., Haas L., Cattaneo R.;
RT "The hemagglutinin of canine distemper virus determines tropism and
cytopathogenicity.";
RL J. Virol. 75:6418-6427(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Ondestepoort;
RA von Messling V.A., Zimmer G., Herzler G., Haas L., Cattaneo R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -I- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF378705; AAK54668.1; -.
DR HSP6; P04849; ISVF.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72898 MW; CC6A104A96BBF8A0 CRC64;

Query Match      87.6%; Score 78; DB 2; Length 662;
Best Local Similarity 88.2%; Pred. No. 0.0001;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
DB 50 PRSDRPVSYTMDIRS 66

RESULT 9
Q6TV24 PRELIMINARY; PRT; 662 AA.
AC Q6TV24;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5804;
RX MEDLINE=22972557; PubMed=14610181;
RX DOI=10.1128/JVI.77.23.12579-12591.2003;
RA von Messling V., Springfield C., Devaux P., Cattaneo R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -I- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -I- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY386315; AAQ96299.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72705 MW; 91F71B34F84DAC0C CRC64;

Query Match      87.6%; Score 78; DB 2; Length 662;
Best Local Similarity 88.2%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
DB 50 PRSDRPVHYIMNRTS 66

RESULT 10
Q6TV24 PRELIMINARY; PRT; 662 AA.
AC Q6TV24;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5804P;
RX MEDLINE=22972557; PubMed=14610181;
RX DOI=10.1128/JVI.77.23.12579-12591.2003;
RA von Messling V., Springfield C., Devaux P., Cattaneo R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -I- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -I- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY386316; AAQ96307.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72733 MW; 43352909C3199DCB CRC64;

Query Match      87.6%; Score 78; DB 2; Length 662;
Best Local Similarity 88.2%; Pred. No. 0.0001;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
DB 50 PRSDRPVHYIMNRTS 66

RESULT 11
O55337 PRELIMINARY; PRT; 266 AA.
AC O55337;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -I- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -I- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026243; AAB88268.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.

```



OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11232;  
RN [1]  
RA SEQUENCE FROM N.A.  
RP Lednický J.A., Meehan T.P., Sarich N.A., Witecki K.E.;  
RT "Full genomic sequence of Canine distemper virus 01-2689.";  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
RC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
CC similarity).  
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
CC family.  
DR EMBL; AY649446; AAT94552.1; -.  
DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; Fusion\_gly; 1.  
KW Envelope protein; Fusion protein.  
KX Envelope protein; Fusion protein.  
SQ SEQUENCE 662 AA; 72565 MW; 656A8757BFD763AA CRC64;

Query Match 71.9%; Score 64; DB 2; Length 662;  
Best Local Similarity 70.6%; Pred. NO. 0.028;  
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PRTSDRPSVYNNRTS 17  
| | | | | : | | | | |  
Db 50 PLTSDRPIHYMNRIS 66

RESULT 14

O6TIB5 PRELIMINARY; PRT; 662 AA.

ID O6TIB5  
AC Q6TIB5;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Fusion protein F.  
DE Canine distemper virus.  
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11232;  
RN [1]  
RA SEQUENCE FROM N.A.  
RP PubMed; 15081610; DOI=10.1016/j.jviromet.2004.02.004;  
RX Lednický J.A., Meehan T.P., Kinzel M.J., Dubach J., Hungerford L.L.,  
SA Sarich N.A., Witecki K.E., Braid M.D., Pedrak C., Houde C.M.;  
RT "Effective primary isolation of wild-type Canine distemper virus in  
MDCK, MVL Lu and Vero cells without nucleotide sequence changes within  
the entire haemagglutinin protein gene and in subgenomic sections of  
the fusion and phosphoprotein genes.";  
RL J. Virol. Methods 118:147-157(2004).  
RN [2].

RA Lednický J.A., Meehan T.P., Sarich N.A.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
RC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
CC similarity).  
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
CC family.  
DR EMBL; AV443350; AAR32272.1; -.  
DR EMBL; AY393984; AAQ96334.1; -.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; Fusion\_gly; 1.  
KW Envelope protein; Fusion protein.  
SQ SEQUENCE 662 AA; 72544 MW; E7B4A3BD7D734B77 CRC64;

Query Match 71.9%; Score 64; DB 2; Length 662;  
Best Local Similarity 70.6%; Pred. NO. 0.028;  
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PRTSDRPSVYNNRTS 17  
| | | | | : | | | | |  
Db 50 PLTSDRPIHYMNRIS 66

## RESULT 15

```

Q9YKL7 ID Q9YKL7 PRELIMINARY; PRT; 662 AA.
AC Q9YKL7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fusion protein.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A75/17;
RX MEDLINE=99139009; PubMed=9971809;
RA Cherpillod P., Beck K., Zurbriggen A., Wittek R.;
RT "Sequence analysis and expression of the attachment and fusion
RL proteins of canine distemper virus wild-type strain A75/17.";
RL J. Virol. 73:2263-2269(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A75/17;
RA Wiederkehr C., Howley P., Zurbriggen A., Wittek R.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -I- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -I- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF112188; AAD18007.1; -.
DR EMBL; AF164967; AAD49702.1; -.
DR HSP; P04849; ISVF.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; I.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72537 MW; 68F992DCBA51F0BA CRC64;

Query Match 71.9%; Score 64; DB 2; Length 662;
Best Local Similarity 70.6%; Pred. No. 0.028;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRISDRPVSYTMNRTS 17
Db 50 PLTSDRPIHYMNRIS 66

```

Search completed: September 28, 2005, 17:55:49  
 Job time : 82.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:40:37 ; Search time 17 Seconds  
(without alignments)  
96.217 Million cell updates/sec

Title: US-10-705-819b-4  
Perfect score: 89  
Sequence: 1 PRTSDRPVSYTMMRTS 17  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	662	1 VGNZCD	cell fusion glycoprotein precursor - canine distemper virus
2	89	100.0	662	2 S21382	cell fusion glycoprotein F1; fusion protein F2
3	42	47.2	374	2 T33844	hypothetical prote
4	42	47.2	403	2 T49645	hypothetical prote
5	42	47.2	632	2 JS0831	alpha-amylase [EC
6	41	46.1	114	2 D84852	hypothetical prote
7	41	46.1	599	1 A64235	cytadherence-acces
8	41	46.1	614	2 T06741	hypothetical prote
9	40	44.9	40	2 T35510	hypothetical prote
10	40	44.9	303	2 A12543	hypothetical prote
11	40	44.9	1206	2 B87247	probable conserved
12	39	43.8	301	2 G70613	probable mmaA4 pro
13	39	43.8	318	2 A11924	heme O synthase [i
14	39	43.8	371	2 T08485	KlcB protein - Ent
15	39	43.8	460	2 H87396	conserved hypotet
16	39	43.8	491	2 C70868	hypothetical prote
17	39	43.8	519	2 S14529	transition protein
18	39	43.8	562	1 S31847	ribosomal protein
19	39	43.8	589	2 AB1151	internalin protein
20	39	43.8	607	2 E96598	protein F2ON2.2 [i
21	39	43.8	699	2 AC0504	probable chitinase
22	39	43.8	825	2 AC0039	glycerol-3-phospha
23	39	43.8	1188	2 T13933	pol polyprotein -
24	39	43.8	1261	2 E59430	PrPLI-associated R
25	38.5	43.3	541	2 AF0547	propionate catabol
26	38	42.7	116	2 S04094	transition protein
27	38	42.7	174	2 B72712	hypothetical prote
28	38	42.7	205	2 B97767	hypothetical prote
29	38	42.7	208	2 A72406	conserved hypotet

30	38	42.7	235	1 WMBEXC	UL56 protein - hum
31	38	42.7	336	2 AF2162	UDP-glucose 4-epim
32	38	42.7	342	1 B71051	conserved hypotet
33	38	42.7	377	2 T49604	hypothetical prote
34	38	42.7	378	2 AC1454	protein gp18 from
35	38	42.7	378	2 AD1090	protein gp18 from
36	38	42.7	515	2 S19487	hypothetical prote
37	38	42.7	539	2 G87272	medium-chain-fatty
38	38	42.7	586	2 T29695	hypothetical prote
39	38	42.7	617	2 T37732	conserved hypotet
40	38	42.7	737	2 T16737	hypothetical prote
41	38	42.7	1072	2 S76888	hypothetical prote
42	38	42.7	1072	2 T18902	hypothetical prote
43	38	42.7	2475	2 S35307	polyprotein pp220
44	37.5	42.1	124	2 D81810	hypothetical prote
45	37.5	42.1	199	2 T49486	hypothetical prote

ALIGNMENTS

RESULT 1

VGNZCD

cell fusion glycoprotein precursor - canine distemper virus

N:Contains: fusion protein F1; fusion protein F2

C:Species: canine distemper virus

C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004

C:Accession: JS0321

R;Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.

Virus Res. 8, 373-386, 1987

A>Title: The nucleotide sequence of the gene encoding the F protein of canine distemper

A;Reference number: JS0321; MUID:88129050; PMID:3433924

A;Accession: JS0321

A;Molecule type: mRNA

A;Residues: 1-662 <BAR>

A;Cross-references: UNIPROT:P12569; GB:M21849; NID:G323241; PIDN:AAA42878.1; PID:G323242

C;Genetics:

A;Gene: F

C;Superfamily: parainfluenza virus cell fusion protein

C;Keywords: glycoprotein; membrane fusion; transmembrane protein

F;1-135/Domain: signal sequence #status predicted <SIG>

F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>

F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>

F;606-629/Domain: transmembrane #status predicted <MEM>

F;62;141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 89; DB 1; Length 662;

Best Local Similarity 100.0%; Pred. No. 4.7e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMMRTS 17

|||||

Db 50 PRTSDRPVSYTMMRTS 66

RESULT 2

S21382

cell fusion protein - canine distemper virus

C;Species: canine distemper virus

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C:Accession: S21382

R;Wild, T.F.; Bernard, A.; Spohner, D.; Villevall, D.; Drillien, R.

submitted to the EMBL Data Library, April 1992

A;Description: Vaccination of mice against canine distemper virus induced encephalitis

A;Reference number: S21382

A;Accession: S21382

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1-662 <WIL>

A;Cross-references: UNIPROT:P12569; EMBL:X65509; NID:G58853; PIDN:CAA46481.1; PID:G58854

C;Superfamily: parainfluenza virus cell fusion protein

Query Match 100.0%; Score 89; DB 2; Length 662;

Best Local Similarity 100.0%; Pred. No. 4.7e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRTSDRPVSYTMNRTS 17  
| | | | | | | | | | | | | | | | | | | | |  
Db 50 PRTSDRPVSYTMNRTS 66

RESULT 3  
T33944  
hypothetical protein Y44E3A.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T33944  
R;Woessner, J.; Graves, T.; Keppler, D.  
submitted to the EMBL Data Library, November 1998  
A;Description: The sequence of C. elegans cosmid Y44E3A.  
A;Reference number: Z21422  
A;Accession: T33944  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-374 <NOE>  
A;Cross-references: UNIPROT:Q9TX9; EMBL:AF106589; PIDN:AAC78229.1; GSPDB:GN00019; CESP:  
A;Experimental source: strain Bristol N2; clone Y44E3A  
C;Genetics:  
A;Gene: CESP:Y44E3A.4  
A;Map position: 1  
A;Introns: 13/3; 68/3; 108/1; 121/3; 175/3; 244/3; 319/1

Query Match 47.2%; Score 42; DB 2; Length 374;  
Best Local Similarity 61.5%; Pred. No. 23;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 DRPVSYTMNRTS 17  
: | | | | | | | | | | | | | | | | | | | | |  
Db 233 NRPMVMNRTS 245

RESULT 4  
T49645  
hypothetical protein B5022.270 [imported] - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C;Accession: T49645  
R;Schulte, U.; Aign, V.; Hobeisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A;Reference number: Z25022  
A;Accession: T49645  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-403 <SCH>  
A;Cross-references: UNIPROT:Q9P5S5; EMBL:AL355932; GSPDB:GN00116; NCSP:B5022.270  
A;Experimental source: BAC clone B5022; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B5022.270  
A;Map position: 6  
A;Introns: 49/2

Query Match 47.2%; Score 42; DB 2; Length 403;  
Best Local Similarity 64.3%; Pred. No. 25;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 TSDRPVSYTMNRT 16  
| : | | | | | | | | | | | | | | | | | | | | |  
Db 80 TGELPVSSSTNLT 93

RESULT 5  
JS0631  
alpha-amylase (EC 3.2.1.1) precursor - Pseudomonas sp.  
N;Alternate names: maltopentaose-forming amylase  
C;Species: Pseudomonas sp.  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004

C;Accession: JS0631  
R;Shida, O.; Takano, T.; Takagi, H.; Kadowaki, K.; Kobayashi, S.  
Biosci. Biotechnol. Biochem. 56, 76-80, 1992  
A;Title: Cloning and nucleotide sequence of the maltopentaose-forming amylase gene from I  
A;Reference number: JS0631; MUID:92257012; PMID:1368136  
A;Accession: JS0631  
A;Molecule type: DNA  
A;Residues: 1-632 <SHI>  
A;Cross-references: UNIPROT:Q52516  
A;Experimental source: strain KO-8940  
A;Note: It is uncertain whether ARG for 1-Met or for 19-Met is the initiation codon  
C;Comment: This enzyme hydrolyzes alpha-1,4-D-glucosidic linkages from the nonreducing er  
C;Function:  
A;Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds  
A;Pathway: glycogen/starch degradation  
C;Superfamily: alpha-amylase; alpha-amylase core homology  
C;Keywords: glycosidase; hydrolase; polysaccharide degradation  
F;1-4/Domain: signal sequence #status predicted <SIG>  
F;45-632/Product: alpha-amylase #status predicted <ALP>  
F;197-332/Domain: alpha-amylase core homology <AMY>

Query Match 47.2%; Score 42; DB 2; Length 632;  
Best Local Similarity 58.3%; Pred. No. 41;  
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 6 RPVSYTMNRTS 17  
: | | | | | | | | | | | | | | | | | | | | |  
Db 105 QPVSYSIDRSRS 116

RESULT 6  
D84852  
hypothetical protein At2g42310 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: D84852  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: D84852  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-114 <STO>  
A;Cross-references: UNIPROT:Q9SLC8; GB:AE002093; NID:g4567303; PIDN:AAD23714.1; GSPDB:GN  
C;Genetics:  
A;Gene: At2g42310  
A;Map position: 2

Query Match 46.1%; Score 41; DB 2; Length 114;  
Best Local Similarity 53.3%; Pred. No. 9.9;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 PRTSDRPVSYTMNRT 15  
| : | | | | | | | | | | | | | | | | | | | | |  
Db 35 PGHEERPKGYFMNRT 49

RESULT 7  
A64235  
cytadherence-accessory protein (hmw3) homolog MG317 - Mycoplasma genitalium  
C;Species: Mycoplasma genitalium  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: A64235  
R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
C.A.; Venter, J.C.  
Science 270, 397-403, 1995  
A;Title: The minimal gene complement of Mycoplasma genitalium.  
A;Reference number: A64200; MUID:96026346; PMID:7569993  
A;Accession: A64235

A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-599 <TIGR>  
A:Cross-references: UNIPROT:Q57081; GB:U39713; GB:L43967; NID:gl046014; PID:gl046018; TI  
A:Experimental source: strain G-37  
C:Genetics:  
A:Genetic code: SGC3  
C:Superfamily: cytochrome-cytochrome protein hmw3

Query Match 46.1%; Score 41; DB 1; Length 599;  
Best Local Similarity 63.6%; Pred. No. 57;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TSDRPVSYTMN 13  
|||:|:|:  
Db 70 TSDNPFSTNN 80

RESULT 8  
T06741  
hypothetical protein F15B8.20 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T06741  
R:Quetier, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Salanoubat, M.; Mewes,  
submitted to the Protein Sequence Database, April 1999  
A:Reference number: Z15794  
A:Accession: T06741  
A:Molecule type: DNA  
A:Residues: 1-614 <QUE>  
A:Cross-references: UNIPROT:Q5VZ3; EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.20  
A:Experimental source: cultivar Columbia; BAC clone F15B8  
C:Genetics:  
A:Gene: ATSP:F15B8.20  
A:Map position: 3  
A:Introns: 198/2; 309/1; 481/3

Query Match 46.1%; Score 41; DB 2; Length 614;  
Best Local Similarity 53.3%; Pred. No. 59;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PRSDRPVSYTMNRT 15  
|||:|:|:  
Db 535 PGHERPKGYFMNRT 549

RESULT 9  
T35510  
hypothetical protein SC6G10.01c - Streptomyces coelicolor (fragment)  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T35510  
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A:Reference number: Z21581  
A:Accession: T35510  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-40 <SEE>  
A:Cross-references: EMBL:AL049497; PIDN:CAB39855.1; GSPDB:GN00070; SCOEDB:SC6G10.01c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC6G10.01c

Query Match 44.9%; Score 40; DB 2; Length 40;  
Best Local Similarity 53.3%; Pred. No. 4.8;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PRSDRPVSYTMNRT 15  
|||:|:|:  
Db 23 PGADAPVPFPMT 37

RESULT 10  
AI2543  
hypothetical protein alr7636 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120b6  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: AI2543  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AI2543  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-303 <KUR>  
A:Cross-references: UNIPROT:Q8ZS76; GB:AF003602; PIDN:BAW7279.1; PID:gl7134721; GSPDB:G  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr7636  
A:Genome: plasmid

Query Match 44.9%; Score 40; DB 2; Length 303;  
Best Local Similarity 63.6%; Pred. No. 41;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 TSDRPVSYTMN 13  
|||:|:|:  
Db 151 TSDXPGTYTFN 161

RESULT 11  
B87247  
probable conserved membrane protein ML2700 [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: B87247  
R:Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: B87247  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1206 <STO>  
A:Cross-references: UNIPROT:Q9CCX9; GB:AL450380; NID:gl3093890; PIDN:CAC32232.1; GSPDB:G  
C:Genetics:  
A:Gene: ML2700

Query Match 44.9%; Score 40; DB 2; Length 1206;  
Best Local Similarity 47.1%; Pred. No. 1.8e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PRSDRPVSYTMNRTS 17  
|||:|:|:  
Db 629 PEVSDRPVESASSRSAS 645

RESULT 12  
G70613  
probable mmaA4 protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: G70613  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: G70613  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-301 <COL>  
A;Cross-references: UNIPROT:P72028; GB:AL123456; NID:g3261722; PIDN:CAB07101.  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: mmaA4

Query Match 43.8%; Score 39; DB 2; Length 301;  
Best Local Similarity 46.7%; Pred. No. 60;  
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RTSDRPVSYTMNR 16  
Db 3 RMAEKPISPKTR 17  
||||:|||||

RESULT 13  
A11924  
heme O synthase [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: A11924  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A;Reference number: AB1807; MUID:21595285; PMID:11759840  
A;Accession: A11924  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-318 <KUR>  
A;Cross-references: UNIPROT:Q8YVA3; GB:BA000019; PIDN:BAB72905.1; PID:gl7130294; GSPDB:G70613  
A;Experimental source: strain PCC 7120  
C;Genetics:  
A;Gene: all0948  
C;Superfamily: heme O synthase

Query Match 43.8%; Score 39; DB 2; Length 318;  
Best Local Similarity 58.3%; Pred. No. 64;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 DRPVSYTMNR 16  
Db 76 DRDIDYEMER 87  
||||:|||||

RESULT 14  
T08485  
k1CB protein - Enterobacter aerogenes plasmid R751  
C;Species: Enterobacter aerogenes  
C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C;Accession: T08485  
R;Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.  
Plasmid 36, 95-111, 1996  
A;Title: Conservation of the genetic switch between replication and transfer genes of IncX1 plasmids  
A;Reference number: Z16434; MUID:97118926; PMID:8954881  
A;Accession: T08485  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-371 <THO>  
A;Cross-references: EMBL:U67194; NID:gl572520; PIDN:AAC64429.1; PID:gl572534  
C;Genetics:  
A;Gene: k1CB  
A;Genome: plasmid R751

Query Match 43.8%; Score 39; DB 2; Length 371;  
Best Local Similarity 70.0%; Pred. No. 75;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 DRPVSYTMNR 14  
Db 336 DRPVYKTR 345  
||||:|||||

RESULT 15  
H87396

conserved hypothetical protein CC1189 [imported] - Caulobacter crescentus  
C;Species: Caulobacter crescentus  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C;Accession: H87396  
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.L.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A;Title: Complete Genome Sequence of Caulobacter crescentus.  
A;Reference number: A87249; MUID:21173698; PMID:11259647  
A;Accession: H87396  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-460 <STO>  
A;Cross-references: UNIPROT:Q9A906; GB:AE005673; NID:gl3422512; PIDN:AAK23172.1; GSPDB:G70613  
C;Genetics:  
A;Gene: CC1189

Query Match 43.8%; Score 39; DB 2; Length 460;  
Best Local Similarity 43.8%; Pred. No. 94;  
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNR 16  
Db 162 PRTKPRACNYALERAR 177  
||||:|||||

Search completed: September 28, 2005, 17:57:03  
Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:41:21 ; Search time 23.5 Seconds  
(without alignments)  
54.001 Million cell updates/sec

Title: US-10-705-819B-25

Perfect score: 88

Sequence: 1 SPDKLLTFIASDTCPLV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	17	4	US-09-890-650-25
2	88	100.0	437	4	US-08-350-841A-1584
3	88	100.0	438	3	US-08-486-099-93
4	88	100.0	438	3	US-08-360-107A-103
5	88	100.0	438	3	US-08-484-223B-93
6	88	100.0	438	3	US-08-919-597-93
7	88	100.0	438	3	US-08-475-668A-93
8	88	100.0	438	3	US-08-485-551A-93
9	88	100.0	438	3	US-08-471-913A-93
10	88	100.0	438	3	US-08-485-264A-93
11	88	100.0	438	3	US-08-474-349A-93
12	88	100.0	438	4	US-08-255-208A-29
13	88	100.0	438	4	US-08-470-896-93
14	88	100.0	438	4	US-08-485-546A-93
15	88	100.0	438	4	US-08-487-266A-93
16	88	100.0	662	1	US-08-234-657-88
17	88	100.0	662	1	US-08-354-138-88
18	88	100.0	662	4	US-09-890-650-27
19	66	75.0	437	4	US-09-350-841A-1591
20	66	75.0	438	3	US-08-486-099-105
21	66	75.0	438	3	US-08-360-107A-115
22	66	75.0	438	3	US-08-484-223B-105
23	66	75.0	438	3	US-08-919-597-105
24	66	75.0	438	3	US-08-475-668A-105
25	66	75.0	438	3	US-08-485-551A-105
26	66	75.0	438	3	US-08-471-913A-105
27	66	75.0	438	3	US-08-485-264A-105

28	66	75.0	438	3	US-08-474-349A-105
29	66	75.0	438	4	US-08-470-896-105
30	66	75.0	438	4	US-08-485-546A-105
31	66	75.0	438	4	US-08-487-266A-105
32	66	75.0	550	1	US-08-279-700-16
33	66	75.0	550	1	US-08-279-700-18
34	66	75.0	550	1	US-08-279-700-20
35	66	75.0	550	1	US-08-279-700-22
36	66	75.0	550	1	US-08-348-891A-5
37	66	75.0	550	2	US-08-905-817-5
38	66	75.0	550	3	US-09-230-944-20
39	66	75.0	550	4	US-09-873-233A-20
40	65	73.9	550	3	US-09-230-944-18
41	65	73.9	550	4	US-09-873-233A-18
42	49	55.7	17	4	US-09-890-650-24
43	47.5	54.0	755	4	US-09-919-497-57
44	44	50.0	59	4	US-09-621-976-5742
45	41.5	47.2	278	1	US-07-921-807B-10

#### ALIGNMENTS

##### RESULT 1

US-09-890-650-25  
; Sequence 25, Application US/09890650  
; Patent No. 6685947  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID CHARLES  
; APPLICANT: SOURAVI, GHOSH  
; APPLICANT: WALKER, JOHN  
; TITLE OF INVENTION: T HELPER CELL EPITOPES  
; FILE REFERENCE: 47-152  
; CURRENT APPLICATION NUMBER: US/09/890,650  
; CURRENT FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 25  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: canine distemper virus  
US-09-890-650-25

Query Match 100.0%; Score 88; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.9e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SPDKLLTFIASDTCPLV	17
Db	1	SPDKLLTFIASDTCPLV	17

##### RESULT 2

US-09-350-841A-1584  
; Sequence 1584, Application US/09350841A  
; Patent No. 6750008  
; GENERAL INFORMATION:  
; APPLICANT: Jeffis, Peter;  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE  
; FILE REFERENCE: 7872-068-999  
; CURRENT APPLICATION NUMBER: US/09/350,841A  
; CURRENT FILING DATE: 1999-07-09  
; NUMBER OF SEQ ID NOS: 1946  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1584  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-350-841A-1584

Query Match 100.0%; Score 88; DB 4; Length 437;  
Best Local Similarity 100.0%; Pred. No. 6.8e-07;

```
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 295 SPDKLTFIASDTCPLV 311

RESULT 3
US-08-486-099-93
; Sequence 93, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; NUMBER OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-486-099-93

Query Match 100.0%; Score 88; DB 3; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 295 SPDKLTFIASDTCPLV 311

RESULT 4
US-08-360-107A-103
; Sequence 103, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
```

```
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107A
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107A-103

Query Match 100.0%; Score 88; DB 3; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 295 SPDKLTFIASDTCPLV 311

RESULT 5
US-08-484-223B-93
; Sequence 93, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
```



Qy 1 SPDKLLTFIASDTCPLV 17

```
Db      295 SPDKLTFIASDTCPLV 311
|||||
RESULT 8
US-08-485-551A-93
; Sequence 93, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-93

Query Match      100.0%; Score 88; DB 3; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SPDKLTFIASDTCPLV 17
|||||
Db      295 SPDKLTFIASDTCPLV 311
|||||

RESULT 10
US-08-485-264A-93
; Sequence 93, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,264A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-021  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 438 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-264A-93

Query Match 100.0%; Score 88; DB 3; Length 438;  
Best Local Similarity 100.0%; Pred. No. 6.8e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17  
|||||  
Db 295 SPDKLLTFIASDTCPLV 311

RESULT 11  
US-08-474-349A-93  
Sequence 93, Application US/08474349A  
Patent No. 633395  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA  
TITLE OF INVENTION: VIRUS TRANSMISSION  
NUMBER OF SEQUENCES: 517  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,349A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 93:

SEQUENCE CHARACTERISTICS:  
LENGTH: 438 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-474-349A-93

Query Match 100.0%; Score 88; DB 3; Length 438;  
Best Local Similarity 100.0%; Pred. No. 6.8e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17  
|||||  
Db 295 SPDKLLTFIASDTCPLV 311

RESULT 12  
US-08-255-208A-29  
Sequence 29, Application US/08255208A  
Patent No. 6440656  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway Jr., Stephen R.  
TITLE OF INVENTION: SYNTHETIC PEPTIDE INHIBITORS OF HIV  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 111  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/255,208A  
FILING DATE: 07-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 438 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-255-208A-29

Query Match 100.0%; Score 88; DB 4; Length 438;  
Best Local Similarity 100.0%; Pred. No. 6.8e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17  
|||||  
Db 295 SPDKLLTFIASDTCPLV 311

RESULT 13

```
US-08-470-896-93
; Sequence 93, Application US/08470896
; Patent No. 6479055
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 896
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-470-896-93

Query Match 100.0%; Score 88; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFFIASDTCPLV 17
|||||
Db 295 SPDKLLTFFIASDTCPLV 311

RESULT 14
US-08-485-546A-93
; Sequence 93, Application US/08485546A
; Patent No. 6518013
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214

US-08-470-896-93
; Sequence 93, Application US/08470896
; Patent No. 6479055
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 93
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-470-896-93

Query Match 100.0%; Score 88; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFFIASDTCPLV 17
|||||
Db 295 SPDKLLTFFIASDTCPLV 311

RESULT 15
US-08-487-266A-93
; Sequence 93, Application US/08487266A
; Patent No. 6824783
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-487-266A-93

Query Match      100.0%; Score 88; DB 4; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 SPDKLLTFIASDTCPLV 17
        |||||
Db      295 SPDKLLTFIASDTCPLV 311

```

Search completed: September 28, 2005, 17:58:45  
Job time : 23.5 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 28, 2005, 17:41:21 ; Search time 23.5 Seconds  
(without alignments)  
54.001 Million cell updates/sec

Title: US-10-705-819B-4  
Perfect score: 89  
Sequence: 1 PRSDRPSVTWNRTRS 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	4	US-09-890-650-4
2	89	100.0	662	1	US-08-224-657-88
3	89	100.0	662	3	US-09-354-138-88
4	89	100.0	662	4	US-09-890-650-27
5	43	48.3	155	4	US-09-252-991A-30279
6	42	47.2	581	4	US-09-252-991A-17734
7	42	47.2	614	3	US-09-017-706-9
8	42	47.2	614	3	US-09-017-706-10
9	42	47.2	614	3	US-09-017-706-11
10	42	47.2	614	3	US-09-017-706-12
11	42	47.2	614	3	US-09-017-706-13
12	42	47.2	614	3	US-09-017-706-14
13	41.5	46.6	248	4	US-09-270-767-42128
14	40	44.9	654	4	US-09-252-991A-19805
15	39	43.8	171	4	US-09-107-532A-7286
16	39	43.8	444	4	US-09-252-991A-23992
17	39	43.8	560	1	US-08-336-618-22
18	39	43.8	1261	3	US-09-080-855-2
19	39	43.8	1261	4	US-09-566-076-2
20	38	42.7	106	4	US-09-270-767-34911
21	38	42.7	106	4	US-09-270-767-50128
22	38	42.7	141	4	US-09-252-991A-19681
23	38	42.7	151	4	US-09-252-991A-16814
24	38	42.7	165	4	US-09-902-540-15191
25	38	42.7	452	4	US-09-328-352-7162
26	38	42.7	466	4	US-09-252-991A-32761
27	38	42.7	514	1	US-08-190-802A-66

28 38 42.7 514 3 US-08-477-346-66 Sequence 66, Appl  
29 38 42.7 514 3 US-08-473-089-66 Sequence 66, Appl  
30 38 42.7 514 4 US-08-487-072A-66 Sequence 66, Appl  
31 38 42.7 577 4 US-09-252-991A-20567 Sequence 20567, A  
32 38 42.7 601 4 US-09-252-991A-21824 Sequence 21824, A  
33 38 42.7 609 4 US-09-949-016-10979 Sequence 10979, A  
34 38 42.7 1157 1 US-07-876-280-30 Sequence 30, Appl  
35 38 42.7 1157 1 US-07-812-180A-2 Sequence 2, Appl  
36 38 42.7 1157 1 US-08-315-468-2 Sequence 2, Appl  
37 38 42.7 1157 3 US-07-941-650A-2 Sequence 2, Appl  
38 37.5 42.1 136 4 US-09-252-991A-20908 Sequence 45516, A  
39 37.5 42.1 141 4 US-09-270-767-45516 Sequence 45516, A  
40 37.5 42.1 325 4 US-09-543-681A-6773 Sequence 6773, Ap  
41 37 41.6 191 4 US-09-252-991A-18626 Sequence 18626, A  
42 37 41.6 278 4 US-09-252-991A-31895 Sequence 31895, A  
43 37 41.6 337 4 US-09-328-352-5833 Sequence 5833, Ap  
44 37 41.6 365 3 US-09-151-592-2 Sequence 2, Appl  
45 37 41.6 365 3 US-09-254-077A-10 Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-09-890-650-4  
; Sequence 4, Application US/09890650  
; Patent No. 6685947  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID CHARLES  
; APPLICANT: SOURAVI, GHOSH  
; APPLICANT: WALKER, JOHN  
; TITLE OF INVENTION: T HELPER CELL EPITOPES  
; FILE REFERENCE: 47-152  
; CURRENT APPLICATION NUMBER: US/09/890,650  
; CURRENT FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: canine distemper virus  
US-09-890-650-4

Query Match 100.0%; Score 89; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRSDRPSVTWNRTRS 17  
Db |||||  
1 PRSDRPSVTWNRTRS 17

RESULT 2  
US-08-224-657-88  
; Sequence 88, Application US/08224657  
; Patent No. 5756102  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: Tartaglia, James  
; APPLICANT: Taylor, Jill  
; TITLE OF INVENTION: FOXPVIRUS - CANINE DISTEMPER VIRUS (CDV)  
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
; TITLE OF INVENTION: RECOMBINANTS  
; NUMBER OF SEQUENCES: 122  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/224,657  
FILING DATE: 06-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2550  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTWS  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 662 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-224-657-88

Query Match 100.0%; Score 89; DB 1; Length 662;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPSVSYTMNRTS 17  
|||||  
DB 50 PRSDRPSVSYTMNRTS 66

RESULT 3  
US-09-354-138-88  
Sequence 88, Application US/09354138  
Patent No. 6309647  
GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
APPLICANT: Tartaglia, James  
APPLICANT: Taylor, Jill  
APPLICANT: Gettig, Russell  
TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
TITLE OF INVENTION: RECOMBINANTS  
NUMBER OF SEQUENCES: 139  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue, 25th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/354,138  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,379  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/416,646  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/224,657  
FILING DATE: 16-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/073,962

FILING DATE: 08-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/776,867  
FILING DATE: 23-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,614  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/938,283  
FILING DATE: 31-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/105,483  
FILING DATE: 12-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/713,967  
FILING DATE: 11-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07,666,056  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2860  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 662 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-09-354-138-88

Query Match 100.0%; Score 89; DB 3; Length 662;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPSVSYTMNRTS 17  
|||||  
DB 50 PRSDRPSVSYTMNRTS 66

RESULT 4  
US-09-890-650-27  
Sequence 27, Application US/09890650  
Patent No. 6685947  
GENERAL INFORMATION:  
APPLICANT: JACKSON, DAVID CHARLES  
APPLICANT: SOURAVI, GHOSH  
APPLICANT: WALKER, JOHN  
TITLE OF INVENTION: T HELPER CELL EPITOPES  
FILE REFERENCE: 47-152  
CURRENT APPLICATION NUMBER: US/09/890,650  
CURRENT FILING DATE: 2001-11-05  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 27  
LENGTH: 662  
TYPE: PRT  
ORGANISM: canine distemper virus  
US-09-890-650-27

Query Match 100.0%; Score 89; DB 4; Length 662;  
Best Local Similarity 100.0%; Pred. No. 4.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPSVSYTMNRTS 17



; GENERAL INFORMATION:  
 ; APPLICANT: ITO, YOSHIFUMI  
 ; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING

```
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-11

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 10
US-09-017-706-12
; Sequence 12, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAIOSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-12

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 11
US-09-017-706-13
; Sequence 13, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAIOSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-13

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 12
US-09-017-706-14
; Sequence 14, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAIOSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-14

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 13
US-09-270-767-42128
; Sequence 42128, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42128
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42128

Query Match      46.6%; Score 41.5; DB 4; Length 248;
Best Local Similarity 52.9%; Pred. No. 51;
Matches 9; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy      1 PRTSDRPVSYTNMRTS 17
Db      90 PNTNDRPISFA-SRTL 105
      :|:|:|:|:|:|

RESULT 14
US-09-252-991A-19805
; Sequence 19805, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-11

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 10
US-09-017-706-12
; Sequence 12, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAIOSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-12

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 11
US-09-017-706-13
; Sequence 13, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAIOSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-13

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 12
US-09-017-706-14
; Sequence 14, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPENTAIOSE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-14

Query Match      47.2%; Score 42; DB 3; Length 614;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy      6 RPVSYTNMRTS 17
Db      87 QPVSYSLDRSRS 98
      :||||:|:|

RESULT 13
US-09-270-767-42128
; Sequence 42128, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42128
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42128

Query Match      46.6%; Score 41.5; DB 4; Length 248;
Best Local Similarity 52.9%; Pred. No. 51;
Matches 9; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy      1 PRTSDRPVSYTNMRTS 17
Db      90 PNTNDRPISFA-SRTL 105
      :|:|:|:|:|:|

RESULT 14
US-09-252-991A-19805
; Sequence 19805, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-19  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 19805  
;; LENGTH: 654  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19805

Query Match 44.9%; Score 40; DB 4; Length 654;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 8; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 PRTSDRPVSYTNRTR 16  
|||  
DB 244 PRPGDRPFGVRAGRTR 259

RESULT 15

US-09-107-532A-7286  
; Sequence 7286, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 7286:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 171 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...171  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7286:  
US-09-107-532A-7286

Query Match 43.8%; Score 39; DB 4; Length 171;  
Best Local Similarity 50.0%; Pred. No. 85;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 2 RTSDRPVSYTN 13  
:::|||||:  
DB 158 KSRDKPVSYRLN 169

Search completed: September 28, 2005, 17:58:44  
Job time : 24.5 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:40:37 ; Search time 17 Seconds  
(without alignments)  
96.217 Million cell updates/sec

Title: US-10-705-819B-25  
Perfect score: 88  
Sequence: 1 SPDKLTFIADTCPLV 17  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	100.0	542	2 JQ2223	cell fusion protei
2	88	100.0	662	1 VGNZCD	cell fusion glycop
3	88	100.0	662	2 S21382	cell fusion protei
4	80	90.9	631	1 A48346	cell fusion glycop
5	80	90.9	631	1 VGNZPD	cell fusion glycop
6	74	84.1	552	2 S47034	cell fusion protei
7	68	77.3	546	1 VGNZRK	cell fusion glycop
8	68	77.3	546	2 S47305	gene F protein - r
9	68	77.3	546	2 S55386	cell fusion protei
10	67	76.1	546	1 VGNZRL	cell fusion glycop
11	67	76.1	546	2 S47300	gene F protein - r
12	67	76.1	636	2 S47399	gene F protein - r
13	66	75.0	282	2 PQ0376	cell fusion glycop
14	66	75.0	282	2 PQ0388	cell fusion glycop
15	66	75.0	550	1 E48556	cell fusion glycop
16	66	75.0	553	1 VGNZMV	cell fusion glycop
17	62	70.5	534	1 JU0274	cell fusion glycop
18	50.5	57.4	1747	2 A45974	collagen alpha 1(X
19	50.5	57.4	1857	2 S31212	collagen alpha 1(X
20	50.5	57.4	1888	2 S78476	collagen alpha 1(X
21	46.5	52.8	3124	2 A40020	collagen alpha 1(X
22	43.5	49.4	26	2 A44036	collagen alpha 1(X
23	43	48.9	1246	2 G64899	nitrate reductase
24	42	47.7	167	2 I48075	aphrodisin precurs
25	41.5	47.2	288	2 E72209	conserved hypothet
26	41.5	47.2	306	2 S09882	hypothetical prote
27	41	46.6	105	2 G71020	hypothetical prote
28	41	46.6	304	2 AD0864	probable membrane
29	41	46.6	378	2 A47300	cell adhesion prot

30	41	46.6	474	2 E85175	ATP-dependent RNA
31	41	46.6	497	2 D84778	probable salt-indu
32	41	46.6	515	1 ISCHSS	protein disulfide-
33	40	45.5	198	2 T27722	hypothetical prote
34	40	45.5	827	2 JC4900	transferred entry
35	40	45.5	1228	2 S46754	hypothetical prote
36	40	45.5	1246	2 G90887	hypothetical prote
37	40	45.5	1246	2 B85730	cryptic nitrate re
38	40	45.5	1305	2 AH3452	cryptic nitrate re
39	40	45.5	1643	2 T07961	myosin heavy chain
40	39.5	44.9	115	2 I52370	transcription migrati
41	39.5	44.9	381	2 S26812	macrophage migrati
42	39	44.3	243	2 A72769	transcription fact
43	39	44.3	283	2 H69179	hypothetical prote
44	39	44.3	293	2 S75355	conserved hypothet
45	39	44.3	311	2 A42982	probable [acyl-car
					rfaS protein - Esc

ALIGNMENTS

RESULT 1

JQ2223  
cell fusion protein F0 precursor - phocine distemper virus  
N:Contains: F1 and F2 chains  
C:Species: phocine distemper virus  
C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C:Accession: JQ2223  
R:Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oervell,  
J. Gen. Virol. 74, 1989-1994, 1993  
A:Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites and  
e virus entity.  
A:Reference number: JQ2223; MUID:93389459; PMID:8376973  
A:Accession: JQ2223  
A:Molecule type: mRNA  
A:Residues: 1-542 <VIS>  
A:Cross-references: UNIPROT:Q7LZY1; GB:I07075  
A:Note: The authors translated the codon ATC for residue 4 as Leu  
C:Comment: This fusion protein F0 is cleaved into F1 and F2 chains.  
C:Genetics:  
A:Gene: F  
C:Superfamily: parainfluenza virus cell fusion protein  
C:Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-15/Domain: signal sequence #status predicted <SIG>  
F;16-542/Product: fusion protein #status predicted <MAT>  
F;16-99/Product: F2 chain #status predicted <F2C>  
F;105-542/Product: F1 chain #status predicted <F1C>  
F;105-135/Region: hydrophobic  
F;486-512/Domain: transmembrane #status predicted <TMM>  
F;21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 88; DB 2; Length 542;  
Best Local Similarity 100.0%; Pred. No. 1.3e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIADTCPLV 17  
|||||

Db 399 SPDKLTFIADTCPLV 415  
|||||

RESULT 2

VGNZCD  
cell fusion glycoprotein precursor - canine distemper virus  
N:Contains: fusion protein F1; fusion protein F2  
C:Species: canine distemper virus  
C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004  
C:Accession: JQ0321  
R:Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.  
Virus Res. 8, 373-386, 1987  
A:Title: The nucleotide sequence of the gene encoding the F protein of canine distemper  
A:Reference number: JQ0321; MUID:88129050; PMID:3433924  
A:Accession: JQ0321  
A:Molecule type: mRNA

```
A;Residues: 1-662 <BAR>
A;Cross-references: UNIPROT:P12569; GB:M21849; NID:G323241; PIDN:AAA42878.1; PID:G323242
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-135/Domain: signal sequence #status predicted <SIG>
F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>
F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>
F;606-629/Domain: transmembrane #status predicted <MEM>
F;62,141,173,179,517/Binding site: carbohydrate (Asn) #status predicted

Query Match 100.0%; Score 88; DB 1; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLFTFIASDTCPLV 17
|||||
DB 519 SPDKLFTFIASDTCPLV 535

RESULT 3
S21382
cell fusion protein - canine distemper virus
C;Species: canine distemper virus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S21382
R;Wild, T.F.; Bernard, A.; Spohner, D.; Villeval, D.; Drillien, R.
submitted to the EMBL Data Library, April 1992
A;Description: Vaccination of mice against canine distemper virus induced encephalitis
A;Reference number: S21382
A;Accession: S21382
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-662 <WIL>
A;Cross-references: UNIPROT:P12569; EMBL:X65509; NID:G58853; PIDN:CAA46481.1; PID:G58854
C;Superfamily: parainfluenza virus cell fusion protein

Query Match 100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLFTFIASDTCPLV 17
|||||
DB 519 SPDKLFTFIASDTCPLV 535

RESULT 4
A48346
cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)
N;Contains: fusion protein F1; fusion protein F2
C;Species: phocine distemper virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: A48346
R;Curran, M.D.; Lu, Y.J.; Rima, B.K.
Arch. Virol. 126, 159-169, 1992
A;Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced amino acid sequence
A;Reference number: A48346; MUID:92396437; PMID:1524494
A;Accession: A48346
A;Molecule type: mRNA
A;Residues: 1-631 <CUR>
A;Cross-references: UNIPROT:P28886
A;Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBIP:113099)
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F;89-106/Domain: transmembrane #status predicted <TM1>
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F;194-219/Domain: transmembrane #status predicted <TM2>
F;575-595/Domain: transmembrane #status predicted <TM3>
F;110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 90.9%; Score 80; DB 1; Length 631;
Best Local Similarity 94.1%; Pred. No. 3e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPDKLFTFIASDTCPLV 17
|||||
DB 488 SPDKLFTFIASDTCPLV 504

RESULT 5
VGNZPD
cell fusion glycoprotein precursor - phocine distemper virus
N;Contains: fusion protein F1; fusion protein F2
C;Species: phocine distemper virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: JQ1368
R;Koevamees, J.; Blixenkrone-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.
J. Gen. Virol. 72, 2959-2966, 1991
A;Title: The nucleotide sequence and deduced amino acid composition of the haemagglutinin
A;Reference number: JQ1368; MUID:92113538; PMID:1765768
A;Accession: JQ1368
A;Molecule type: genomic RNA
A;Residues: 1-631 <KOV>
A;Cross-references: UNIPROT:P28886
C;Genetics:
A;Gene: F
C;Superfamily: parainfluenza virus cell fusion protein
C;Keywords: glycoprotein; membrane fusion; transmembrane protein
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>
F;89-106/Domain: transmembrane #status predicted <TM1>
F;189-193/Region: cleavage processing #status predicted
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>
F;194-212/Domain: transmembrane #status predicted <TM2>
F;575-595/Domain: transmembrane #status predicted <TM3>
F;110,142,146,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.9%; Score 80; DB 1; Length 631;
Best Local Similarity 94.1%; Pred. No. 3e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SPDKLFTFIASDTCPLV 17
|||||
DB 488 SPDKLFTFIASDTCPLV 504

RESULT 6
S47034
cell fusion protein precursor - porpoise morbillivirus
N;Alternate names: F protein
C;Species: porpoise morbillivirus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S47034
R;Bolt, G.; Gottschalk, E.; Blixenkrone-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.; Earl
submitted to the EMBL Data Library, July 1994
A;Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbilli-
A;Reference number: S47034
A;Accession: S47034
A;Molecule type: mRNA
A;Residues: 1-552 <BOL>
A;Cross-references: UNIPROT:Q66147; EMBL:X80757; NID:G520639; PIDN:CAA56731.1; PID:G52064
A;Experimental source: isolate Ulster 88
A;Note: the source is designated as Cetacean morbillivirus
C;Superfamily: parainfluenza virus cell fusion protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-552/Product: fusion protein #status predicted <MAT>

Query Match 84.1%; Score 74; DB 2; Length 552;
Best Local Similarity 81.2%; Pred. No. 0.00026;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLLTFIASDTCPLV 17
|||||
```

submitted to the EMBL Data Library, September 1994

A;Description: The nucleotide sequence of fusion protein gene of the Peete des petits r to each virus.

A;Reference number: S55386

A;Accession: S55386

A;Molecule type: DNA

A;Residues: 1-546 <MEY>

A;Cross-references: UNIPROT:Q84926; EMBL:Z37017; NID:G854372; PIDN:CAA85451.1; PID:G8543

A;Experimental source: strain 75/1; cell line vero

C;Genetics:

A;Gene: F.

C;Superfamily: parainfluenza virus cell fusion protein

C;Keywords: membrane fusion

Query Match 77.3%; Score 68; DB 2; Length 546;  
Best Local Similarity 81.2%; Pred. No. 0.0025; 2; Indels 0; Gaps 0;  
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PDKLLTFTIASDTCPVLV 17  
||||| ||||| |||  
Db 404 PDKLLTFTIASDKCPV 419

RESULT 10  
VGNZRL  
cell fusion glycoprotein precursor - rinderpest virus (strain L)  
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C;Species: rinderpest virus  
C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
C;Accession: A28921  
R;Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.  
Virology 164, 523-530, 1988  
A;Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of the  
A;Reference number: A28921; MUID:88219541; PMID:3285575  
A;Accession: A28921  
A;Molecule type: mRNA  
A;Residues: 1-546 <TSU>  
A;Cross-references: UNIPROT:P10864; GB:M20870; NID:G333898; PIDN:AAA47399.1; PID:G333899  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: Glycoprotein; membrane fusion; transmembrane protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-104/Product: cell fusion glycoprotein F2 #status predicted <FG2>  
F;105-546/Product: cell fusion glycoprotein F1 #status predicted <FG1>  
F;109-133/Product: transmembrane #status predicted <TN1>  
F;485-513/Domain: transmembrane #status predicted <TN2>  
F;25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 76.1%; Score 67; DB 1; Length 546;  
Best Local Similarity 68.8%; Pred. No. 0.0037; 1; Indels 0; Gaps 0;  
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDKLLTFTIASDTCPVLV 17  
||||| ||||| |||  
Db 404 PDKLLTFTIASDKCPV 419

RESULT 11  
S47300  
gene F protein - rinderpest virus  
C;Species: rinderpest virus  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S47300; PQ0865  
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A;Description: The complete nucleotide sequence of the fusion protein gene of the vacci  
A;Reference number: S47299  
A;Accession: S47300  
A;Molecule type: DNA  
A;Residues: 1-546 <EVA>  
A;Cross-references: UNIPROT:P41360; EMBL:Z31656; NID:G535406; PIDN:CAA83482.1; PID:G5354  
R;Chamberlain, R.W.; Wanmay, H.M.; Hockley, E.; Shalla, M.S.; Goatley, L.; Knowles, N.J.

J. Gen. Virol. 74, 2775-2780, 1993  
A;Title: Evidence for different lineages of rinderpest virus reflecting their geographic  
A;Reference number: PQ0865; MUID:94103786; PMID:8277286  
A;Accession: PQ0865  
A;Molecule type: mRNA  
A;Residues: 86-191 <CHA>  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 76.1%; Score 67; DB 2; Length 546;  
Best Local Similarity 68.8%; Pred. No. 0.0037;  
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLTFIADSDTCPLV 17  
||||:||||:||||:  
Db 404 PDKILTYIAADQCPVV 419

RESULT 12  
S47299  
Gene F protein - rinderpest virus  
C;Species: rinderpest virus  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S47299  
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T  
submitted to the EMBL Data Library, March 1994  
A;Description: The complete nucleotide sequence of the fusion protein gene of the vacci  
A;Reference number: S47299  
A;Accession: S47299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-636 <EVA>  
A;Cross-references: UNIPROT:Q86486; EMBL:Z31655; NID:q535391; PIDN:CAA83481.1; PID:g5353  
C;Superfamily: parainfluenza virus cell fusion protein

Query Match 76.1%; Score 67; DB 2; Length 636;  
Best Local Similarity 68.8%; Pred. No. 0.0043;  
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLTFIADSDTCPLV 17  
||||:||||:||||:  
Db 494 PDKILTYIAADQCPVV 509

RESULT 13  
PQ0376  
cell fusion glycoprotein - measles virus (strain TT) (fragment)  
C;Species: measles virus  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: PQ0376  
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparison  
A;Reference number: PQ0374; MUID:92300360; PMID:1607874  
A;Accession: PQ0376  
A;Molecule type: Genomic RNA  
A;Residues: 1-282 <SCH>  
A;Cross-references: UNIPROT:Q83629; UNIPROT:Q9IC36; UNIPROT:Q88973; UNIP  
PROT:P90330; UNIPROT:Q9QEW7; UNIPROT:Q9WMK4; UNIPROT:Q83525; UNIPROT:Q83518; UNIPROT:Q89  
3521; UNIPROT:Q83530; UNIPROT:Q91248; UNIPROT:Q91Q22; UNIPROT:Q9QEW8; UNIPROT:Q04244  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion

Query Match 75.0%; Score 66; DB 2; Length 282;  
Best Local Similarity 68.8%; Pred. No. 0.0029;  
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLTFIADSDTCPLV 17  
||||:||||:||||:  
Db 408 PDKILTYIAADQCPVV 423

Search completed: September 28, 2005, 17:57:06  
Job time : 17 secs

Db 140 PDKILTYIAADHCPVV 155

RESULT 14  
PQ0388  
cell fusion glycoprotein - measles virus (strain Schwarz vaccine) (fragment)  
C;Species: measles virus  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: PQ0388  
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparison  
A;Reference number: PQ0374; MUID:92300360; PMID:1607874  
A;Accession: PQ0388  
A;Molecule type: Genomic RNA  
A;Residues: 1-282 <SCH>  
A;Cross-references: UNIPROT:Q83525; UNIPROT:Q83530  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion

Query Match 75.0%; Score 66; DB 2; Length 282;  
Best Local Similarity 68.8%; Pred. No. 0.0029;  
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLTFIADSDTCPLV 17  
||||:||||:||||:  
Db 140 PDKILTYIAADHCPVV 155

RESULT 15  
E48556  
cell fusion glycoprotein precursor - measles virus (strain AIK-C)  
C;Species: measles virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: E48556  
R;Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.  
Virus Genes 7, 67-81, 1993  
A;Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C  
A;Reference number: E48556; MUID:93227570; PMID:8470368  
A;Accession: E48556  
A;Molecule type: Genomic RNA  
A;Residues: 1-550 <MOR>  
A;Cross-references: UNIPROT:P35973; GB:S58435; NID:G299460; PIDN:BA826145.1; PID:G299465  
A;Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBIP:129272)  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <PF2>  
F;108-550/Product: cell fusion glycoprotein F1 #status predicted <PF1>  
F;113-138/Region: hydrophobic  
F;495-514/Domain: transmembrane #status predicted <TMW>  
F;6,29,61,67/Binding site: carbohydrate (Asn) #status predicted

Query Match 75.0%; Score 66; DB 1; Length 550;  
Best Local Similarity 68.8%; Pred. No. 0.0054;  
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLTFIADSDTCPLV 17  
||||:||||:||||:  
Db 408 PDKILTYIAADHCPVV 423



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:39:51 ; Search time 80.5 Seconds  
(without alignments)  
108.141 Million cell updates/sec

Title: US-10-705-819B-25  
Perfect score: 88  
Sequence: 1 SPDKLLTFFIASDTCPLV 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	528	2 Q9YJW9	Q9YJW9 canine dist
2	88	100.0	530	2 Q8QV06	Q8QV06 canine dist
3	88	100.0	542	2 Q7LZV1	Q7LZV1 phocine dis
4	88	100.0	662	1 VGLF_CDVO	P12569 canine dist
5	88	100.0	662	2 Q89327	Q89327 canine dist
6	88	100.0	662	2 Q91K33	Q91K33 canine dist
7	88	100.0	662	2 Q68PH9	Q68PH9 canine dist
8	88	100.0	662	2 Q6SYT0	Q6SYT0 canine dist
9	88	100.0	662	2 Q6TIB5	Q6TIB5 canine dist
10	88	100.0	662	2 Q6TV24	Q6TV24 canine dist
11	88	100.0	662	2 Q6TV32	Q6TV32 canine dist
12	88	100.0	662	2 Q9DX22	Q9DX22 canine dist
13	88	100.0	662	2 Q9YKL7	Q9YKL7 canine dist
14	80	90.9	631	1 VGLF_PHODV	P28866 phocine dis
15	74	84.1	552	2 Q56852	Q56852 dolphin mor
16	74	84.1	552	2 Q66147	Q66147 cetacean mo
17	74	84.1	552	2 Q66409	Q66409 dolphin mor
18	68	77.3	546	1 VGLF_RINDK	P12574 rinderpest
19	68	77.3	546	1 VGLF_RINDR	P41356 rinderpest
20	68	77.3	546	2 Q91H55	Q91H55 rinderpest
21	68	77.3	546	2 Q84926	Q84926 peste-des-p
22	68	77.3	553	2 Q9IC36	Q9IC36 measles vir
23	67	76.1	546	1 VGLF_RINDB	P41360 rinderpest
24	67	76.1	546	1 VGLF_RINDL	P10864 rinderpest
25	67	76.1	546	2 Q6Q439	Q6Q439 peste-des-p
26	67	76.1	550	2 Q8V049	Q8V049 measles vir
27	67	76.1	553	2 Q91248	Q91248 measles vir
28	67	76.1	579	2 Q9PMU4	Q9PMU4 measles vir
29	67	76.1	636	2 Q86486	Q86486 rinderpest
30	66	75.0	138	2 Q11758	Q11758 measles vir
31	66	75.0	162	2 Q11739	Q11739 measles vir

## ALIGNMENTS

## RESULT 1

Q9YJW9 ID Q9YJW9 PRELIMINARY; PRT; 528 AA.  
AC Q9YJW9;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DB Fusion protein precursor (Fragment)  
OS Canine distemper virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11232;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Liermann H., Harder T.C., Loechelt M., Baumgaertner W., Moennig V.,  
RA Haas L.;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Harder T.C., von Messling V., Oervell C., Moennig V., Haas L.;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
CC -|- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By similarity)  
CC -|- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein family.  
CC EMBL; AJ007711; CAA07617.1; -  
DR HSSP; P04849; ISVF.  
DR GO; GO:0019031; C: viral envelope; IEA.  
DR GO; GO:0006948; P: viral-induced cell-cell fusion; IEA.  
DR InterPro; IPR000776; Fusion\_gly.  
DR Pfam; PF00523; Fusion\_gly; 1.  
KW Envelope protein; Fusion protein; Signal.  
FT NON TER 1  
FT SIGNAL <1 10 potential.  
FT CHAIN 11 >528 fusion protein.  
FT NON TER 528 528  
SQ SEQUENCE 528 AA; 57613 MW; 146C8CFBF6EF6516 CRC64;

Query Match 100.0%; Score 88; DB 2; Length 528;  
Best Local Similarity 100.0%; Pred. No. 8.5e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFFIASDTCPLV 17  
|||  
DB 392 SPDKLLTFFIASDTCPLV 408

## RESULT 2

Q8QV06 ID Q8QV06 PRELIMINARY; PRT; 530 AA.  
AC Q8QV06;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Andersen M.K.; AAL83966.1; -.
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF355188; AAL83966.1; -.
DR HSP; P04849; ISVF.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0006948; P: viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
FT NON TER 1
SQ SEQUENCE 530 AA; 57984 MW; 8F7173C247AF233D CRC64;

Query Match 100.0%; Score 88; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 8.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17
Db 387 SPDKLLTFIASDTCPLV 403

RESULT 3
Q7LZY1
ID 07LZY1 PRELIMINARY; PRT; 542 AA.
AC Q7LZY1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Cell fusion protein F0 precursor.
OS Phocine distemper virus (PDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11240;
RN [1]
RP SEQUENCE FROM N.A.
RA Visser I.K.G.; van der Heijden R.W.J.; van de Bildt M.W.G.;
RA Kenter M.J.H.; Oervell C.; Osterhaus A.D.M.E.;
RT "Fusion protein gene nucleotide sequence similarities, shared
RT antigenic sites and phylogenetic analysis suggest that phocid
RT distemper virus type 2 and canine distemper virus belong to the same
RT virus entity.";
RL J. Gen. Virol. 74:1989-1994(1993).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR PIR; JQ2223; JQ2223.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0006948; P: viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 542 AA; 59286 MW; 57F1F81C6B54AA72 CRC64;

Query Match 100.0%; Score 88; DB 2; Length 542;
Best Local Similarity 100.0%; Pred. No. 8.7e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17
|||||
```

```
Db 399 SPDKLLTFIASDTCPLV 415

RESULT 4
VGLF CDVO STANDARD; PRT; 662 AA.
ID VGLF CDVO
AC P12569; Q65991;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;
DE Fusion glycoprotein F1].
GN Name=F;
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88129050; PubMed=3433924; DOI=10.1016/0168-1702(87)90009-8;
RA Barrett T.; Clarke D.K.; Evans S.A.; Rima B.K.;
RT "The nucleotide sequence of the gene encoding the F protein of canine
RT distemper virus: a comparison of the deduced amino acid sequence with
RT other paramyxoviruses.";
RL Virus Res. 8:373-386(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=932727696; PubMed=8470428; DOI=10.1016/0264-410X(93)90285-6;
RA Wild T.F.; Bernard A.; Spehner D.; Villevall D.; Drillean R.;
RT "Vaccination of mice against canine distemper virus-induced
RT encephalitis with vaccinia virus recombinants encoding measles or
RT canine distemper virus antigens.";
RL Vaccine 11:438-444(1993).
CC -1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked.
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21849; AAA42878.1; -.
DR EMBL; X65509; CAA46481.1; -.
DR PIR; JS0321; VGNZCD.
DR PIR; S21382; S21382.
DR HSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 ?
FT CHAIN ? 662 Fusion glycoprotein F0.
FT CHAIN ? 224 Fusion glycoprotein F2.
FT CHAIN 225 662 Fusion glycoprotein F1.
FT TRANSMEM 606 629 Potential.
FT DISULFID 180 307 Linkage between F2 and F1 (Potential).
FT CARBOHYD 62 62 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 141 141 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 173 173 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 179 179 N-linked (GlcNAc...) (Potential).
FT CONFLICT 3 3 R -> K (in Ref. 2).
FT CONFLICT 140 140 D -> N (in Ref. 2).
FT CONFLICT 152 152 N -> S (in Ref. 2).
FT CONFLICT 171 171 I -> M (in Ref. 2).
FT CONFLICT 174 174 A -> V (in Ref. 2).
FT CONFLICT 662 662 L -> H (in Ref. 2).
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;
```

```

Query Match      100.0%; Score 88; DB 1; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 519 SPDKLTFIASDTCPLV 535

RESULT 5
ID O89327 PRELIMINARY; PRT; 662 AA.
AC O89327;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93174978; PubMed=8438593;
RA Sidhu M.S.; Husar W.; Cook S.D.; Dowling P.C.; Udem S.A.;
RT "Canine distemper terminal and intergenic non-protein coding
RL nucleotide sequences: completion of the entire CDV genome sequence.";
RN Virology 193:66-72(1993).
RN [2]
RN SEQUENCE FROM N.A.
RA Sidhu M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF014953; AAC26994.1; -.
DR HSP; P04849; ISVF.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin-like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72951 MW; 80B144C6B9801898 CRC64;

Query Match      100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 519 SPDKLTFIASDTCPLV 535

RESULT 6
ID Q91KN3 PRELIMINARY; PRT; 662 AA.
AC Q91KN3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Onderstepoort;
RX MEDLINE=21306344; PubMed=11413309;
RX DOI=10.1128/JVI.75.14.6418-6427.2001;
RA von Messling V.; Zimmer G.; Herrier G.; Haas L.; Cattaneo R.;

```

```

RT "The hemagglutinin of canine distemper virus determines tropism and
RT cytopathogenicity.";
RL J. Virol. 75:6418-6427(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Onderstepoort;
RA von Messling V.A.; Zimmer G.; Herrier G.; Haas L.; Cattaneo R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF378705; AAK54668.1; -.
DR HSP; P04849; ISVF.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72898 MW; CC6A104A96BBF8A0 CRC64;

Query Match      100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 519 SPDKLTFIASDTCPLV 535

RESULT 7
ID Q68PH9 PRELIMINARY; PRT; 662 AA.
AC Q68PH9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fusion protein.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RN SEQUENCE FROM N.A.
RA Lednický J.A.; Meehan T.P.; Sarich N.A.; Witecki K.E.;
RT "Full genomic sequence of Canine distemper virus 01-2689.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY649446; AAT94552.1; -.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72565 MW; 656A8757BFD763AA CRC64;

Query Match      100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 519 SPDKLTFIASDTCPLV 535

RESULT 8
ID Q6SYO0 PRELIMINARY; PRT; 662 AA.
AC Q6SYO0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fusion protein F.

```

```

OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Lednický J.A., Meehan T.P., Sarich N.A.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;
RA Lednický J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,
RA Sarich N.A., Wittecki K.E., Braid M.D., Pedrak C., Houde C.M.;
RT "Effective primary isolation of wild-type Canine distemper virus in
RT MDCK, Mv1 Lu and Vero cells without nucleotide sequence changes within
RT the entire haemagglutinin protein gene and in subgenomic sections of
RT the fusion and phosphoprotein genes.";
RL J. Virol. Methods 118:147-157(2004).
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY445077; AAR16539.1; -.
DR EMBL; AY542312; AAS48411.1; -.
DR EMBL; AY466011; AAR30102.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72736 MW; 2DD1B45291D3D315 CRC64;

Query Match 100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17
DB 519 SPDKLLTFIASDTCPLV 535

RESULT 9
ID Q6T1B5 PRELIMINARY; PRT; 662 AA.
AC Q6T1B5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 28, Last sequence update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;
RA Lednický J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,
RA Sarich N.A., Wittecki K.E., Braid M.D., Pedrak C., Houde C.M.;
RT "Effective primary isolation of wild-type Canine distemper virus in
RT MDCK, Mv1 Lu and Vero cells without nucleotide sequence changes within
RT the entire haemagglutinin protein gene and in subgenomic sections of
RT the fusion and phosphoprotein genes.";
RL J. Virol. Methods 118:147-157(2004).
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY443350; AAR32272.1; -.

```

```

DR EMBL; AY395984; AAQ96334.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72544 MW; E7B4A3BD7D734B77 CRC64;

Query Match 100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17
DB 519 SPDKLLTFIASDTCPLV 535

RESULT 10
ID Q6TV24 PRELIMINARY; PRT; 662 AA.
AC Q6TV24;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5804P;
RX MEDLINE=22972557; PubMed=14610181;
RX DOI=10.1128/JVI.77.23.12579-12591.2003;
RA Von Messling V., Springfeld C., Devaux P., Cattaneo R.;
RT "A ferret model of canine distemper virus virulence and
RT immunosuppression.";
RL J. Virol. 77:12579-12591(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5804P;
RA von Messling V., Springfeld C., Devaux P., Cattaneo R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY386316; AAQ96307.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72733 MW; 43352909C3199DCB CRC64;

Query Match 100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17
DB 519 SPDKLLTFIASDTCPLV 535

RESULT 11
ID Q6TV32 PRELIMINARY; PRT; 662 AA.
AC Q6TV32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Fusion protein F.

```

OS Canine distemper virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11232;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5804;  
RX MEDLINE=22972557; PubMed=14610181;  
RA DOI=10.1128/JVI.77.23.12579-12591.2003;  
RA Von Messling V., Springfield C., Devaux P., Cattaneo R.;  
RT "A ferret model of canine distemper virus virulence and  
RT immunosuppression.";  
RL J. Virol. 77:12579-12591(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5804;  
RA Von Messling V., Springfield C., Devaux P., Cattaneo R.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
CC similarity)  
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
CC family.  
DR EMBL; AY386315; AAQ96299.1; -.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
DR InterPro; IPR000776; Fusion\_gly.  
DR InterPro; IPR009050; Globin\_like.  
DR Pfam; PF00523; Fusion\_gly; I.  
KW Envelope protein; Fusion protein.  
SQ SEQUENCE 662 AA; 72705 MW; 91F71B34F84DAC0C CRC64;

Query Match 100.0%; Score 88; DB 2; Length 662;  
Best Local Similarity 100.0%; Pred. NO. 1.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SPDKLTFIASDTCPLV 17  
Db 519 SPDKLTFIASDTCPLV 535  
|||||  
|||||

RESULT 12  
Oy Q9DXZ2 PRELIMINARY; PRT; 662 AA.  
AC Q9DXZ2;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Fusion protein F.  
OS Canine distemper virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11232;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20493096; PubMed=11044118;  
RX DOI=10.1128/JVI.74.22.10737-10744.2000;  
RA Gassen U., Collins F.M., Duprex W.P., Rima B.K.;  
RT "Establishment of a rescue system for canine distemper virus.";  
RL J. Virol. 74:10737-10744(2000).  
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
CC similarity)  
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
CC family.  
DR EMBL; AF305419; AAG30919.1; -.  
DR HSSP; P04849; ISVF.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
DR InterPro; IPR000776; Fusion\_gly.  
DR InterPro; IPR009050; Globin\_like.  
DR Pfam; PF00523; Fusion\_gly; I.  
KW Envelope protein; Fusion protein.  
SQ SEQUENCE 662 AA; 72988 MW; 9C5C1398C8AE7B4C CRC64;

Query Match 100.0%; Score 88; DB 2; Length 662;  
Best Local Similarity 100.0%; Pred. NO. 1.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SPDKLTFIASDTCPLV 17  
Db 519 SPDKLTFIASDTCPLV 535  
|||||  
|||||

RESULT 13  
Oy Q9YKL7 PRELIMINARY; PRT; 662 AA.  
AC Q9YKL7;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Fusion protein.  
OS Canine distemper virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11232;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A75/17;  
RX MEDLINE=99139009; PubMed=9971809;  
RA Cherpillod P., Beck K., Zurbriggen A., Wittek R.;  
RT "Sequence analysis and expression of the attachment and fusion  
RT proteins of canine distemper virus wild-type strain A75/17.";  
RL J. Virol. 73:2263-2269(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A75/17;  
RA Wiederkehr C., Howley P., Zurbriggen A., Wittek R.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
CC similarity)  
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
CC family.  
DR EMBL; AF112188; AAD18007.1; -.  
DR EMBL; AF164967; AAD49702.1; -.  
DR HSSP; P04849; ISVF.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
DR InterPro; IPR000776; Fusion\_gly.  
DR InterPro; IPR009050; Globin\_like.  
DR Pfam; PF00523; Fusion\_gly; I.  
KW Envelope protein; Fusion protein.  
SQ SEQUENCE 662 AA; 72537 MW; 68F992DCBA51F0BA CRC64;

Query Match 100.0%; Score 88; DB 2; Length 662;  
Best Local Similarity 100.0%; Pred. NO. 1.1e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SPDKLTFIASDTCPLV 17  
Db 519 SPDKLTFIASDTCPLV 535  
|||||  
|||||

RESULT 14  
Oy VGLF\_PHODV STANDARD; PRT; 631 AA.  
ID VGLF\_PHODV  
AC P28886;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
DE Fusion glycoprotein F1].  
GN Name=F;  
OS Phocine distemper virus (PDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=11240;  
RN [1]

```
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate DX88-4A;
RX MEDLINE=92113538; PubMed=1765768;
RA Koevamees J., Blixenkron-Moeller M., Sharma B., Oervell C.,
RA Norrby E.;
RT "The nucleotide sequence and deduced amino acid composition of the
RT haemagglutinin and fusion proteins of the morbillivirus phocid
RT distemper virus.";
RL J. Gen. Virol. 72:2959-2966(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Uster/88;
RX MEDLINE=92398437; PubMed=1524494;
RA Curran M.D., Lu Y.J., Rima B.K.;
RT "The fusion protein gene of phocine distemper virus: nucleotide and
RT deduced amino acid sequences and a comparison of morbillivirus fusion
RT proteins.";
RL Arch. Virol. 126:159-169(1992).
RN [3]
RP SEQUENCE OF 95-631 FROM N.A.
RC STRAIN=Uster/88;
RX MEDLINE=91089508; PubMed=2264246;
RA Curran M.D., Loan D.O., Rima B.K., Kennedy S.;
RT "Nucleotide sequence analysis of phocine distemper virus reveals its
RT distinctness from canine distemper virus.";
RL Vet. Rec. 127:430-431(1990).
CC -1- FUNCTION: This protein directs fusion of viral and cellular
CC membranes.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked.
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10371; BAA01206.1; -.
DR PIR; A48346; A48346.
DR PIR; JQ1368; VGNZPD.
DR HSSP; P04849; ISVF.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 ?
FT CHAIN ? 631 Fusion glycoprotein F0.
FT CHAIN ? 188 Fusion glycoprotein F2.
FT CHAIN 194 631 Fusion glycoprotein F1.
FT DISULFID 149 276 Linkage between F2 and F1 (Potential).
FT TRANSMEM 89 106 Potential.
FT TRANSMEM 194 212 Potential.
FT TRANSMEM 575 595 Potential.
FT CARBOHYD 110 110 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 142 142 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 148 148 N-linked (GlcNAc. . .) (Potential).
FT CONFLICT 63 63 I -> V (in Ref. 2).
SQ SEQUENCE 631 AA; 68873 MW; D1FC87CDD426E9B8 CRC64;

Query Match 90.9%; Score 80; DB 1; Length 631;
Best Local Similarity 94.1%; Pred. No. 2.7e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SPDKLTFIASDTCPLV 17
Db 488 SPDKLTFIASDTCPLV 504
|||||
|||||

RESULT 15
O56852 PRELIMINARY; PRT; 552 AA.
ID O56852
```

```
AC O56852;
DT -01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F protein.
OS Dolphin morbillivirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=37131;
RN [1]
RP SEQUENCE FROM N.A.
RA Soethout E., Harder T.C., Osterhaus A.D.M.E.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Harder T.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases. (By
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
CC EMBL; AJ224704; CAA12077.1; -.
DR HSSP; P04849; ISVF.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 552 AA; 59770 MW; 80FF6A0F25AF3589 CRC64;

Query Match 84.1%; Score 74; DB 2; Length 552;
Best Local Similarity 81.2%; Pred. No. 0.00028;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDKLTFIASDTCPLV 17
Db 410 PDKLTFVAADKCPV 425
|||||
|||||

Search completed: September 28, 2005, 17:55:52
Job time : 80.5 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:37:01 ; Search time 91.25 Seconds  
(without alignments)  
72.054 Million cell updates/sec

Title: US-10-705-819B-25

Perfect score: 88

Sequence: 1 SPDKLTFIADTCPLV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	17	3 AAB08100	Aab08100 Amino aci
2	88	100.0	17	8 ADK00512	Adk00512 Immunogen
3	88	100.0	17	8 ADJ84586	Adj84586 T-helper
4	88	100.0	438	6 ABO10249	Abol10249 Canine di
5	88	100.0	662	2 AAR83304	Aar83304 Canine di
6	88	100.0	662	3 AAB08102	Aab08102 Amino aci
7	88	100.0	662	4 AAM50124	Aam50124 Canine di
8	88	100.0	662	5 AAM47653	Aam47653 Canine di
9	88	100.0	662	7 ADM66120	Adm66120 vaccinia
10	86	97.7	439	4 AAU14029	Aau14029 Peptide s
11	66	75.0	436	4 AAU14096	Aau14096 Peptide s
12	66	75.0	438	6 ABO10261	Abol10261 Measles v
13	66	75.0	550	2 AAR34540	Aar34540 F protein
14	66	75.0	550	2 AAR42396	Aar42396 Chicago l
15	66	75.0	550	2 AAR42394	Aar42394 Moraten h
16	66	75.0	550	2 AAR42397	Aar42397 Consensus
17	66	75.0	550	2 AAR42395	Aar42395 San Diego
18	66	75.0	550	2 AAW94760	Aaw94760 Mutant me
19	66	75.0	550	4 AAB73722	Aab73722 Measles v
20	66	75.0	550	4 AAB73721	Aab73721 Measles v
21	66	75.0	553	7 ADF86346	Adf86346 Measles v
22	65	73.9	550	2 AAW94759	Aaw94759 Mutant me
23	53	60.2	53	3 AAB27160	Aab27160 Measles v
24	49	55.7	17	3 AAB08099	Aab08099 Amino aci
25	49	55.7	17	8 ADK00511	Adk00511 Immunogen

26	49	55.7	17	8 ADJ84585	Adj84585 T-helper
27	47.5	54.0	755	5 AAU84267	Aau84267 Human end
28	47.5	54.0	795	4 AAB27229	Abab27229 Human EXM
29	47.5	54.0	1780	7 ADE15980	Ade15980 G-coupled
30	47.5	54.0	1780	8 ADL93919	Adl93919 Human G-C
31	47	53.4	372	8 ADN24706	Adn24706 Bacterial
32	47	53.4	372	8 ADN24486	Adn24486 Bacterial
33	47	53.4	377	8 ADN21949	Adn21949 Bacterial
34	47	53.4	377	8 ADN21731	Adn21731 Bacterial
35	43.5	49.4	1899	8 ADJ75667	Adj75667 Marker ge
36	43.5	49.4	3063	5 ABB90762	Abb90762 Human Tum
37	43.5	49.4	3063	6 ABU54469	Abu54469 Human Tum
38	43.5	49.4	3063	6 ABR47415	Abr47415 Breast ca
39	43.5	49.4	3063	6 ABR47416	Abr47416 Breast ca
40	43.5	49.4	3063	8 ADJ75666	Adj75666 Marker ge
41	43.5	49.4	3067	8 ADJ76366	Adj76366 Marker ge
42	43.5	49.4	3118	4 AAU27790	Aau27790 Human ful
43	43	48.9	59	4 ABG29085	Abg29085 Novel hum
44	43	48.9	526	6 ABM67142	Abm67142 Photorhab
45	43	48.9	662	4 ABG17704	Abg17704 Novel hum

## ALIGNMENTS

## RESULT 1

AAB08100  
ID AAB08100 standard; peptide; 17 AA.  
XX  
AC AAB08100;  
XX  
DT 04-DEC-2000 (first entry)  
XX  
DE Amino acid sequence of a helper T cell epitope from CDV.  
XX  
KW T helper cell epitope; CDV; immune response; canine vaccine.  
XX  
OS Canine distemper virus.  
XX  
PN WO200046390-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 07-FEB-2000; 2000WO-AU0000070.  
XX  
PR 05-FEB-1999; 99AU-00008533.  
PR 04-AUG-1999; 99AU-00002013.  
XX  
(UYME ) UNIV MELBOURNE.  
PA (CSLC-) CSL LTD.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PI Jackson DC, Souravi G, Walker J;  
XX WPI; 2000-532904/48.  
XX  
PT Novel T helper cell epitopes derived from canine distemper virus useful for preparation of canine vaccines.  
XX  
PS Claim 1; Page 28; 54pp; English.  
XX  
CC AAB08076-B08101 represent T helper cell epitopes, derived from canine distemper virus (CDV). Compositions comprising these T cell helper epitopes are useful for inducing an immune response in an animal. The CC epitopes are useful as components of animal, in particular, canine CC vaccines, either simply as synthetic peptide based antigens CC additions to vaccines containing more complex antigens  
SQ Sequence 17 AA;

Query Match 100.0%; Score 88; DB 3; Length 17;

```
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17
   |||||
Db 1 SPDKLLTFIASDTCPLV 17

RESULT 2
ID ADK00512
XX ADK00512 standard; peptide; 17 AA.
AC ADK00512;
XX
XX
XX 06-MAY-2004 (first entry)
XX
XX Immunogenic lipopeptide of the invention #48.
XX
XX T helper cell epitope; B cell epitope; Antibacterial; Antiulcer;
KW Antiinfectivity; vaccine; antibody.
KW
XX Synthetic.
XX
XX WO2004014956-A1.
XX
XX 19-FEB-2004.
XX
XX 12-AUG-2003; 2003WO-AU001018.
XX
XX 12-AUG-2002; 2002US-0402838P.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
XX Jackson D, Zeng W;
XX
XX WPI; 2004-238735/22.
XX
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,
PT useful for eliciting immune response against group A Streptococcus
PT antigen.
XX
XX Disclosure; SEQ ID NO 48; 194pp; English.
XX
XX The present invention relates to a lipopeptide comprising polypeptide
CC conjugated to lipid moieties, where polypeptide contains amino acid
CC sequence of T helper cell epitope and B cell epitope, where amino acid
CC sequences are different, and internal lysine residues or internal lysine
CC analog residues for covalent attachment of each of lipid moieties through
CC aegri; amino group or terminal side chain group of lysine or lysine
CC analog. The peptides are useful in eliciting the production of antibody
CC against an antigenic B cell epitope in a subject, and are useful for
CC antibody production, synthetic vaccine production, diagnostic method
CC employing antibodies and antibody ligands and immunotherapy for
CC veterinary and human medicine. The method efficiently elicits the
CC production of antibody against antigenic B cell epitope. The present
CC sequence represents a novel immunogenic lipopeptide comprising T helper
CC and B cell epitopes.
XX
XX Sequence 17 AA;
SQ
Query Match 100.0%; Score 88; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17
   |||||
Db 1 SPDKLLTFIASDTCPLV 17

RESULT 3
ADJ84586
ID ADJ84586 standard; peptide; 17 AA.
```

```
XX
AC ADJ84586;
XX
DT 06-MAY-2004 (first entry)
XX
DE T-helper epitope peptide SEQ ID NO:44.
XX
KW lipopeptide; lipid moiety; T helper cell epitope; Th epitope;
KW cytotoxic T cell epitope; CTL epitope; immune response; immunisation;
KW influenza; hepatitis C virus; Listeria monocytogenes; vaccine; cancer;
KW cytostatic; virucide.
XX
OS Canine distemper virus.
OS Synthetic.
XX
XX WO2004014957-A1.
XX
XX 19-FEB-2004.
XX
XX 12-AUG-2003; 2003WO-AU001019.
XX
XX 12-AUG-2002; 2002US-0403328P.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
XX Jackson D, Zeng W;
XX
XX WPI; 2004-238736/22.
XX
XX New lipopeptide comprising a polypeptide comprising an amino acid
PT sequence of a T helper cell and cytotoxic T cell epitope, useful for
PT preparing a composition for treating or preventing cancer, or hepatitis C
PT virus or influenza virus.
XX
XX Disclosure; SEQ ID NO 44; 166pp; English.
XX
XX The present invention describes a lipopeptide comprising a polypeptide
CC conjugated to one or more lipid moieties, where the polypeptide comprises
CC an amino acid sequence that comprises: (a) the amino acid sequence of a T
CC helper cell (Th) epitope and the amino acid sequence of a cytotoxic T
CC cell (CTL) epitope, where the amino acid sequences are different; and (b)
CC one or more internal lysine residues or internal lysine analogue residues
CC for covalent attachment of each of the lipid moieties via the epsilon-
CC amino group or terminal side-chain group of the lysine or lysine
CC analogue. Each of the one or more lipid moieties is covalently attached
CC to an epsilon-amino group of the one or more internal lysine residues or
CC to a terminal side-chain group of the one or more internal lysine
CC analogue residues. Also described: (1) producing a lipopeptides; (2) a
CC composition comprising the lipopeptide and an excipient or diluent; (3)
CC eliciting an immune response in a subject; (4) immunising a subject
CC against influenza or hepatitis C virus or Listeria monocytogenes; (5) a
CC vaccine against an influenza virus or hepatitis C virus or Listeria
CC monocytogenes comprising the lipopeptides, where the CTL epitope is from
CC an influenza virus or hepatitis C virus or Listeria monocytogenes protein
CC ; (6) treating or preventing cancer; and (7) a vaccine against cancer
CC comprising the lipopeptide, where the CTL epitope is a tumour-specific
CC CTL epitope. The lipopeptide has cytostatic and virucide activities, and
CC can be used in vaccine. The lipopeptide is useful for preparing a
CC composition for treating or preventing cancer, or hepatitis C virus or
CC influenza virus. The present sequence is used in the exemplification of
CC the present invention.
XX
XX Sequence 17 AA;
SQ
Query Match 100.0%; Score 88; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17
   |||||
Db 1 SPDKLLTFIASDTCPLV 17
```



```

RESULT 4
ABO10249
ID ABO10249 standard; protein; 438 AA.
XX AC ABO10249;
XX DT 19-AUG-2003 (first entry)
XX DE Canine distemper virus F1 fusion glycoprotein.
XX KW HIV; DP107; DP178; glycoprotein 41; antiviral; virucide; EBV;
XX KW Epstein-Barr virus infection; heptad repeat motif.
XX OS Canine distemper virus strain Onderstepoort.
XX PN US6518013-B1.
XX PD 11-FEB-2003.
XX PF 07-JUN-1995; 95US-00485546.
XX PR 07-JUN-1993; 93US-00073028.
XX PR 07-JUN-1994; 94US-00255208.
XX PR 20-DEC-1994; 94US-00360107.
XX PA (TRIM-) TRIMERIS INC.
XX PI Barney SO, Lambert DM, Petteway SR;
XX WPI; 2003-465599/44.
XX DT Inhibiting transmission of Epstein-Barr virus to a cell, by contacting
XX PT the cell with a peptide consisting of a region of Epstein-Barr virus
XX PS protein.
XX PS Example; Fig 23; 716pp; English.
XX CC The invention relates to inhibiting (M) transmission of an Epstein-Barr
XX CC virus to a cell, comprising contacting the cell with an effective
XX CC concentration of a peptide consisting of a region of 16-39 consecutive
XX CC amino acids of an Epstein-Barr virus protein for an effective period of
XX CC time, where the region is recognised by one or more of ALLMOT15,
XX CC 107x178x4 or PLZIP sequence search motifs, the peptide further comprises
XX CC an amino terminal X, and a carboxy terminal Z in which X comprises an
XX CC amino group, acetyl group, 9-fluorenylmethoxy-carbonyl group, hydrophobic
XX CC group or macromolecular carrier group, and Z comprises a carboxyl group,
XX CC amido group, hydrophobic group, or macromolecular carrier group, and
XX CC fusion of the virus to the cell is inhibited. The peptides were
XX CC identified by analysing the structure/motifs present in the HIV-1
XX CC glycoprotein 41 anti-HIV peptides DP107 and DP178. These heptad repeat
XX CC motif containing peptides were used to design the motifs cited above,
XX CC which in turn were used to analyse proteins from other pathogenic
XX CC organisms and HIV isolates, looking for DP107/178 structural analogues.
XX CC The method is useful for inhibiting transmission of Epstein-Barr virus to
XX CC a cell and Epstein-Barr virus infection. The present sequence is a
XX CC protein from a pathogenic organism analysed for regions analogous to
XX CC DP107 or DP178
XX SQ Sequence 438 AA;

Query Match 100.0%; Score 88; DB 6; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17
Db 295 SPDKLLTFIASDTCPLV 311
|||||
|||||

RESULT 5
AAR83304
ID AAR83304 standard; protein; 662 AA.
XX AC AAR83304;
XX DT 25-MAR-2003 (revised)
XX DT 29-MAY-1996 (first entry)
XX DE Canine distemper virus fusion glycoprotein.
XX KW Avian influenza virus; rabies; canine distemper virus; CDV; measles;
XX KW antibody; vaccine; therapy; TROVAC.
XX OS Synthetic.
XX PN W09527780-A1.
XX PD 19-OCT-1995.
XX PF 06-APR-1995; 95WO-US004394.
XX PR 06-APR-1994; 94US-00224657.
XX PR 05-APR-1995; 95US-00416646.
XX PA (VIRO-) VIROGENETICS CORP.
XX PI Paoletti E, Tartaglia J, Taylor J, Gettig R;
XX WPI; 1995-366385/47.
XX N-PSDB; AAT00521.
XX DT New modified recombinant viruses - contg. a canine distemper virus
XX PT antigen or measles virus M or N antigen in a non-essential region.
XX PS Example 15; Fig 15; 194pp; English.
XX CC This sequence represents the H6 promoted canine distemper virus (CDV)
XX CC fusion glycoprotein sequence. The DNA encoding this sequence is contained
XX CC in the plasmid pSCDVHA. The encoding sequence was removed from this
XX CC plasmid and inserted into the C5 locus of another plasmid. The vector can
XX CC then be used to introduce the exogenous DNA into a fowlpox virus. This
XX CC procedure can be performed using exogenous DNA encoding a measles virus M
XX CC or N antigen. The recombinant viruses can then be used to induce an
XX CC antigenic or immunological response. The expression products of these
XX CC recombinants, and the antibodies produced can be used in binding assays
XX CC to determine the presence of CDV or measles virus in a sample. The
XX CC attenuated virulence of the viruses reduces the possibility of a runaway
XX CC infection due to the vaccination, in a vaccinated individual. The viruses
XX CC also reduce the transmission from vaccinated to unvaccinated individuals,
XX CC and reduces environmental contamination. (Updated on 25-MAR-2003 to
XX CC correct PR field.)
XX SQ Sequence 662 AA;

Query Match 100.0%; Score 88; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIASDTCPLV 17
Db 519 SPDKLLTFIASDTCPLV 535
|||||
|||||

RESULT 6
AAR808102
ID AAR808102 standard; protein; 662 AA.
XX AC AAR808102;
XX DT 04-DEC-2000 (first entry)
XX DE Amino acid sequence of a fusion protein of canine distemper virus.
XX KW T helper cell epitope; CDV; immune response; canine vaccine.
XX OS Canine distemper virus.

```

```

XX PN WO200046390-A1.
XX XX
XX PD 10-AUG-2000.
XX XX
XX PF 07-FEB-2000; 2000WO-AU000070.
XX XX
XX PR 05-FEB-1999; 99AU-00008533.
XX PR 04-AUG-1999; 99AU-00002013.
XX XX
XX PA (UYME ) UNIV MELBOURNE.
XX PA (CSLC-) CSL LTD.
XX PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX XX
XX PI Jackson DC, Souravi G, Walker J;
XX XX
XX DR WPI; 2000-532904/48.
XX XX
XX PT Novel T helper cell epitopes derived from canine distemper virus useful
XX PT for preparation of canine vaccines.
XX XX
XX PS Example 1; Fig 1; 54pp; English.
XX XX
XX CC The present sequence represents a fusion protein of canine distemper
XX CC virus (CDV). The protein was used to identify peptides AAB08076-B08101,
XX CC which are T helper cell epitopes. Compositions comprising these T cell
XX CC helper epitopes are useful for inducing an immune response in an animal.
XX CC The epitopes are useful as components of animal, in particular, canine
XX CC vaccines, either simply as synthetic peptide based vaccines and as
XX CC additions to vaccines containing more complex antigens
XX XX
XX SQ Sequence 662 AA;
XX XX
XX Query Match 100.0%; Score 88; DB 3; Length 662;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-06;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 1 SPDKLLTFIASDTCPLV 17
XX Db |||||
XX 519 SPDKLLTFIASDTCPLV 535
XX XX
XX RESULT 7
XX AAM50124
XX ID AAM50124 standard; protein; 662 AA.
XX AC AAM50124;
XX XX
XX DT 21-DEC-2001 (first entry)
XX XX
XX DE Canine distemper virus recombinant fusion protein PCDVF662.
XX DE CDV; fusion protein; PCDVF662; antigen; immune status;
XX KW vaccination status; dog.
XX KW Canine distemper virus.
XX OS WO200166568-A2.
XX XX
XX PN 13-SEP-2001.
XX XX
XX PF 07-MAR-2001; 2001WO-US0007251.
XX XX
XX PR 09-MAR-2000; 2000US-00521738.
XX XX
XX PA (HESK-) HESKA CORP.
XX PA (COLS ) UNIV COLORADO STATE RES FOUND.
XX XX
XX PI Jensen WA, Lappin MR, Rosen DK, Andrews JS;
XX XX
XX DR WPI; 2001-639000/73.
XX XX
XX N-PSDB; AAH27071.
XX XX
XX PT Determining immune status or vaccination status of an animal to e.g.
XX PT calicivirus comprises using a recombinant viral antigen.
XX XX
XX PS Claim 4; Page 129-131; 132pp; English.
XX XX
XX CC The present sequence is that of canine distemper virus fusion protein
XX CC recombinant antigen, PCDVF662. The recombinant antigen, fused to an N-
XX CC terminal His tag, was produced in Escherichia coli cells transformed by a
XX CC recombinant vector comprising nucleic acid nCDVF1986 (see AAH27071).
XX CC PCDVF662 is an example of a recombinant infectious agent antigen that can
XX CC be used in the method of the invention to determine the immune status of
XX CC an animal. The method involves contacting a biological specimen of an
XX CC animal (cat, dog or horse) with a recombinant antigen, and detecting the
XX CC presence or absence of a complex between the recombinant antigen and an
XX CC antibody present in the sample. The method determines whether the animal
XX CC is protected against disease or should be vaccinated. Recombinant
XX CC antigens (see AAM50107-24), nucleic acids encoding them (see AAH27054-
XX CC 71), methods of producing them, and assay methods are provided
XX XX
XX SQ Sequence 662 AA;
XX XX
XX Query Match 100.0%; Score 88; DB 4; Length 662;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-06;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 1 SPDKLLTFIASDTCPLV 17
XX Db |||||
XX 519 SPDKLLTFIASDTCPLV 535
XX XX
XX RESULT 8
XX AAM47653
XX ID AAM47653 standard; protein; 662 AA.
XX AC AAM47653;
XX XX
XX DT 21-FEB-2002 (first entry)
XX XX
XX DE Canine distemper virus, CDV, F.
XX DE Virucide; vaccine; virus; virulence; canine distemper virus; CDV;
XX KW measles; dog.
XX KW Canine distemper virus.
XX OS US6309647-B1.
XX PN 30-OCT-2001.
XX XX
XX PF 15-JUL-1999; 99US-00354138.
XX XX
XX PR 15-JUL-1999; 99US-00354138.
XX XX
XX PA (AVET ) AVENTIS PASTEUR.
XX XX
XX PI Paolletti E, Tartaglia J, Taylor J, Gettig R;
XX XX
XX DR WPI; 2002-040232/05.
XX DR N-PSDB; ABI98922.
XX XX
XX PT Novel virus, useful for inducing immune response in dog against CDV,
XX PT comprises the modified recombinant virus having attenuated virulence
XX PT comprising exogenous DNA sequences encoding antigens of canine distemper
XX PT virus (CDV) or measles virus.
XX XX
XX PS Example 15; Fig 15; 147pp; English.
XX XX
XX CC The present invention relates to modified recombinant viruses, comprising
XX CC inactivated virus-encoded genetic functions so that the viruses have
XX CC attenuated virulence, yet retained efficiency. The viruses can contain
XX CC DNA encoding a canine distemper virus (CDV) antigen or measles M or N

```

CC antigen. The recombinant viruses are useful for inducing an antigenic or  
 CC immunological response in a dog or other carnivore against CDV. The  
 CC present sequence was used in an example from the present invention  
 XX  
 SQ Sequence 662 AA;  
 Query Match 100.0%; Score 88; DB 5; Length 662;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SPDKLTFIASDTCPLV 17  
 |||||  
 DB 519 SPDKLTFIASDTCPLV 535  
 RESULT 9  
 ADM66120  
 ID ADM66120 standard; protein; 662 AA.  
 AC ADM66120;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DE Vaccinia virus H6-promoted CDV Fusion (F) protein #1.  
 XX  
 KW Canine distemper virus; NVAC; TROVAC; ALVAC; rabies glycoprotein G;  
 XX rabies; vaccine.  
 XX  
 OS Canine distemper virus; Onderstepoort strain.  
 XX  
 PN US2003082204-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PP 13-SEP-2001; 2001US-00951061.  
 XX  
 PR 20-NOV-1990; 90US-00621614.  
 PR 07-MAR-1991; 91US-00666056.  
 PR 11-JUN-1991; 91US-00713967.  
 PR 22-OCT-1991; 91US-00776867.  
 PR 06-MAR-1992; 92US-00847951.  
 PR 31-AUG-1992; 92US-00938283.  
 PR 08-JUN-1993; 93US-00073962.  
 PR 12-AUG-1993; 93US-00105483.  
 PR 06-APR-1994; 94US-00224657.  
 PR 15-JUL-1999; 99US-00354138.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Paoletti E, Tartaglia J, Taylor J, Gettig R;  
 XX  
 DR WPI; 2003-567445/53.  
 DR N-PSDB; ADM66154.  
 XX  
 XX New recombinant viruses comprising exogenous DNA encoding rabies  
 PT glycoprotein G useful for eliciting protective immunity against rabies  
 PT virus in a carnivore.  
 XX  
 XX Example 15; SEQ ID NO 86; 93pp; English.  
 XX  
 XX The invention relates to recombinant vaccinia and canarypox viruses  
 CC comprising exogenous DNA encoding rabies glycoprotein G in a nonessential  
 CC region of the virus genome. Also included are a recombinant vaccinia virus  
 CC (comprising exogenous DNA encoding rabies glycoprotein G in a  
 CC nonessential region of the virus genome, where at least one open reading  
 CC frame (ORF) selected from J2, B13, B14R, A26L, 156R, C7L-KL, and 14L is  
 CC deleted from the virus), a recombinant canarypox virus (produced by  
 CC attenuation through multiple serial passages on chick embryo fibroblasts,  
 CC subjecting a master seed from to successive plaque purifications under  
 CC agar and amplifying a plaque clone through multiple additional passages,  
 CC where the virus contains exogenous DNA encoding rabies glycoprotein G in  
 CC a nonessential region of the virus genome), inducing an antigenic or  
 CC immunological response in a carnivore against rabies virus (by

CC administering to the dog, cat or other carnivore a composition comprising  
 CC the virus above in a mixture with a carrier), expressing a gene product  
 CC in a cell cultured in vitro by introducing into the cell a virus of the  
 CC invention. One or more (optionally all) ORFs selected from a thymidine  
 CC kinase gene, a haemorrhagic region, an A type inclusion body, a  
 CC haemagglutinin gene, a host range region, and a ribonucleotide reductase  
 CC large subunit gene, may also be deleted. The attenuated Vaccinia virus is  
 CC termed a NVAC virus. The attenuated canarypox virus is termed ALVAC  
 CC recombinant virus. Also included are attenuated fowl pox viruses termed  
 CC TROVAC. The recombinant viruses are useful as vaccines for protecting a  
 CC dog, cat or other carnivore against rabies. The modified recombinant  
 CC viruses are effective as vaccines and are safer than some other  
 CC recombinant viruses due to the deletion of genes affecting virulence that  
 CC are not essential for virus growth in tissue culture. The present  
 CC sequence is a Canine distemper virus protein expressed from an attenuated  
 CC virus of the invention.  
 XX  
 SQ Sequence 662 AA;  
 Query Match 100.0%; Score 88; DB 7; Length 662;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SPDKLTFIASDTCPLV 17  
 |||||  
 DB 519 SPDKLTFIASDTCPLV 535  
 RESULT 10  
 AAU14029  
 ID AAU14029 standard; peptide; 439 AA.  
 XX  
 AC AAU14029;  
 XX  
 DT 21-NOV-2001 (first entry)  
 XX  
 DE Peptide sequence from canine distemper virus fusion glycoprotein Fl.  
 XX  
 KW Anti-retroviral; DP178-like; DP107-like; fusion glycoprotein Fl;  
 XX antitumorogenic; antiviral; HIV transmission.  
 XX  
 OS Canine distemper virus strain Onderstepoort.  
 XX  
 PN WO200151673-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 05-JUL-2000; 2000WO-US035727.  
 XX  
 PR 09-JUL-1999; 99US-00350841.  
 XX  
 PA (TRIM-) TRIMERIS INC.  
 XX  
 PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;  
 XX  
 DR WPI; 2001-442157/47.  
 XX  
 XX Identifying a compound that inhibits the formation of or disrupts a  
 PT DP107/DP178 complex, especially compounds with antitumorogenic, antiviral  
 PT or intracellular modulatory activity, by detecting the formation of a  
 PT DP107/DP178 complex.  
 XX  
 XX Disclosure; Fig 23; 259pp; English.  
 XX  
 XX The present invention relates to peptides which exhibit anti-retroviral  
 CC activity. The peptides of the invention (AAU12559-AAU14009) comprise  
 CC DP178-like and DP107-like peptides. The DP178 peptide corresponds to  
 CC amino acids 639-673 of the transmembrane protein gp41 from human  
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide  
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention  
 CC also relates to a method of identifying compounds that inhibit the  
 CC formation of or disrupts a DP107/DP178 complex. The method comprises  
 CC detecting the formation of a DP107/DP178 complex, both in the presence or

CC absence of a test compound, in a reaction mixture containing DP107 and  
 CC DP178 peptides. The method is useful for identifying compounds, including  
 CC small molecule compounds, which may themselves exhibit antifusogenic,  
 CC antiviral or intracellular modulatory activity. The DP178-like/DP107-like  
 CC peptides are useful to inhibit human and non-human retroviral,  
 CC particularly HIV, transmission to uninfected cells. The present sequence  
 CC represents a peptide sequence from canine distemper virus strain  
 CC Onderstepoort fusion glycoprotein F1

XX  
 SQ Sequence 439 AA;

Query Match 97.7%; Score 86; DB 4; Length 439;  
 Best Local Similarity 94.1%; Pred. No. 3e-06;  
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLTFIASDTCPLV 17  
 |||||:|||||  
 Db 296 SPDKLTFIASDTCPLV 312

RESULT 11  
 AAU14096  
 ID AAU14096 standard; peptide; 436 AA.  
 XX  
 AC AAU14096;  
 XX  
 DT 11-SEP-2003 (revised)  
 DT 21-NOV-2001 (first entry)  
 XX  
 DE Peptide sequence from Measles virus fusion glycoprotein F1.  
 XX  
 KW Anti-retroviral; DP178-like; DP107-like; fusion glycoprotein F1;  
 KW antifusogenic; antiviral; HIV transmission.  
 XX  
 OS Measles virus; strain Edmonston.  
 XX  
 PN WC200151673-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 05-JUL-2000; 2000WO-US035727.  
 XX  
 PR 09-JUL-1999; 99US-00350841.  
 XX  
 PA (TRIM-) TRIMERIS INC.  
 XX  
 PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;  
 XX  
 DR WPI; 2001-442157/47.  
 XX  
 PS Identifying a compound that inhibits the formation of or disrupts a  
 PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral  
 PT or intracellular modulatory activity, by detecting the formation of a  
 PT DP107/DP178 complex.  
 XX  
 PS Disclosure; Fig 34; 259pp; English.  
 XX  
 CC The present invention relates to peptides which exhibit anti-retroviral  
 CC activity. The peptides of the invention (AAU12559-AAU14009) comprise  
 CC DP178-like and DP107-like peptides. The DP178 peptide corresponds to  
 CC amino acids 639-673 of the transmembrane protein gp41 from human  
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide  
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention  
 CC also relates to a method of identifying compounds that inhibit the  
 CC formation of or disrupts a DP107/DP178 complex. The method comprises  
 CC detecting the formation of a DP107/DP178 complex, both in the presence or  
 CC absence of a test compound, in a reaction mixture containing DP107 and  
 CC DP178 peptides. The method is useful for identifying compounds, including  
 CC small molecule compounds, which may themselves exhibit antifusogenic,  
 CC antiviral or intracellular modulatory activity. The DP178-like/DP107-like  
 CC peptides are useful to inhibit human and non-human retroviral,  
 CC particularly HIV, transmission to uninfected cells. The present sequence  
 CC represents a peptide sequence from Measles virus strain Edmonston fusion

CC glycoprotein F1. (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 436 AA;

Query Match 75.0%; Score 66; DB 4; Length 436;  
 Best Local Similarity 68.8%; Pred. No. 0.0091;  
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDKLTFIASDTCPLV 17  
 |||||:|||||  
 Db 294 PDKLTFIAADHCPV 309

RESULT 12  
 ABO10261  
 ID ABO10261 standard; protein; 438 AA.  
 XX  
 AC ABO10261;  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 19-AUG-2003 (first entry)  
 XX  
 DE Measles virus fusion glycoprotein F1.  
 XX  
 KW HIV; DP107; DP178; Glycoprotein 41; antiviral; virucide; EBV;  
 KW Epstein-Barr virus infection; heptad repeat motif.  
 XX  
 OS Measles virus; strain Edmonston.  
 XX  
 PN US6518013-B1.  
 XX  
 PD 11-FEB-2003.  
 XX  
 PF 07-JUN-1995; 95US-00485546.  
 XX  
 PR 07-JUN-1993; 93US-00073028.  
 PR 07-JUN-1994; 94US-00255208.  
 PR 20-DEC-1994; 94US-00360107.  
 XX  
 PA (TRIM-) TRIMERIS INC.  
 XX  
 PI Barney SO, Lambert DM, Petteway SR;  
 XX  
 DR WPI; 2003-465599/44.  
 XX  
 PT Inhibiting transmission of Epstein-Barr virus to a cell, by contacting  
 PT the cell with a peptide consisting of a region of Epstein-Barr virus  
 PT protein.  
 XX  
 PS Example; Fig 34; 716pp; English.  
 CC  
 CC The invention relates to inhibiting (M) transmission of an Epstein-Barr  
 CC virus to a cell, comprising contacting the cell with an effective  
 CC concentration of a peptide consisting of a region of 16-39 consecutive  
 CC amino acids of an Epstein-Barr virus protein for an effective period of  
 CC time, where the region is recognised by one or more of ALLMOT15,  
 CC 107x178x4 or PLZIP sequence search motifs, the peptide further comprises  
 CC an amino terminal X, and a carboxy terminal Z in which X comprises an  
 CC amino group, acetyl group, 9-fluorenylmethoxy-carbonyl group, hydrophobic  
 CC group or macromolecular carrier group, and Z comprises a carboxyl group,  
 CC amide group, hydrophobic group, or macromolecular carrier group, and  
 CC fusion of the virus to the cell is inhibited. The peptides were  
 CC identified by analysing the structure/motifs present in the HIV-1  
 CC glycoprotein 41 anti-HIV peptides DP107 and DP178. These heptad repeat  
 CC motif containing peptides were used to design the motifs cited above,  
 CC which in turn were used to analyse proteins from other pathogenic  
 CC organisms and HIV isolates, looking for DP107/178 structural analogues.  
 CC The method is useful for inhibiting transmission of Epstein-Barr virus to  
 CC a cell and Epstein-Barr virus infection. The present sequence is a  
 CC protein from a pathogenic organism analysed for regions analogous to  
 CC DP107 or DP178. (Updated on 23-Oct-2003 to standardise OS field)

XX Sequence 438 AA;

```

Query Match      75.0%; Score 66; DB 6; Length 438;
Best Local Similarity 68.8%; Pred. No. 0.0091;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 PDKLLTFIASDTCPLV 17
DB      296 PDKILTYIAADHCPV 311

RESULT 13
AAR34540
ID AAR34540 standard; protein; 550 AA.
XX
AC AAR34540;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 19-AUG-1993 (first entry)
XX
DE F protein of attenuated measles virus strain AIK-C.
XX
KW paramyxoviridae; RNA virus; attenuation; vaccine.
XX
OS Measles virus.
XX
PN EP540135-A2.
XX
PD 05-MAY-1993.
XX
PF 10-MAR-1992; 92EP-00302004.
XX
PR 14-OCT-1991; 91JP-00293625.
XX
PA (KITA) KITASATO INST.
XX
PI Sasaki K, Mori T, Makino S;
XX
DR WPI; 1993-145503/18.
XX
DR N-PSDB; AAQ40480.
XX
PT New attenuated measles vaccine virus strain - retains high immunogenicity
PT with reduced pyrogenicity and having no neurological complications.
XX
PS Disclosure; Page 14-18; 47pp; English.
XX
CC The sequence of the entire genome of the attenuated measles virus strain
CC AIK-C has been determined and contains 6 open reading frames. The
CC sequence of the F protein was deduced from the fourth ORF. The virus
CC retains high immunogenicity with reduced pyrogenicity and no neurological
CC complications. See AAR34537-R34541 and AAR39592. (Updated on 25-MAR-2003
CC to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 550 AA;

Query Match      75.0%; Score 66; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.012;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 PDKLLTFIASDTCPLV 17
DB      408 PDKILTYIAADHCPV 423

RESULT 14
AAR42396
ID AAR42396 standard; protein; 550 AA.
XX
AC AAR42396;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 13-MAY-1994 (first entry)
XX
DE Moraten haemagglutinin fusion protein.
XX
KW Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;
KW vaccine; infection; consensus polypeptide.
XX
OS Measles virus strain Moraten; f.
XX
PN WO9321325-A1.
XX
PD 28-OCT-1993.
XX
PF 08-APR-1993; 93WO-US003209.
XX

```

```

XX Chicago 1 haemagglutinin fusion protein.
DE
XX Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;
KW vaccine; infection; consensus polypeptide.
XX
OS Measles virus; strain Chicago 1 f.
XX
PN WO9321325-A1.
XX
PD 28-OCT-1993.
XX
PF 08-APR-1993; 93WO-US003209.
XX
PR 08-APR-1992; 92US-00866033.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Rota JS, Bellini WJ;
XX
DR WPI; 1993-351735/44.
XX
DR N-PSDB; AAQ51097.
XX
PT Haemagglutinin and fusion glycoprotein of several wild-type measles
PT strains - used to construct vaccines for measles infection.
XX
PS Disclosure; Page 68-71; 119pp; English.
XX
CC HA (AAQ51088-94) and fusion glycoprotein (AAQ51095-97) sequences of
CC several wild-type measles strains are given. Shared amino acid variations
CC in wild-type measles glycoproteins are identified in five wild-type
CC measles viruses. A consensus polypeptide, the amino acid sequence of
CC which reflects variation common to more than one wild-type strain, is
CC the basis for constructing live attenuated vaccines, or recombinant
CC vaccines to replace older, less efficacious vaccines. Immunological
CC reagents useful in differentiating wild-type measles strains from other
CC known strains can also be produced. (Updated on 25-MAR-2003 to correct PN
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 550 AA;

Query Match      75.0%; Score 66; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.012;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 PDKLLTFIASDTCPLV 17
DB      408 PDKILTYIAADHCPV 423

RESULT 15
AAR42394
ID AAR42394 standard; protein; 550 AA.
XX
AC AAR42394;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 13-MAY-1994 (first entry)
XX
DE Moraten haemagglutinin fusion protein.
XX
KW Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;
KW vaccine; infection; consensus polypeptide.
XX
OS Measles virus strain Moraten; f.
XX
PN WO9321325-A1.
XX
PD 28-OCT-1993.
XX
PF 08-APR-1993; 93WO-US003209.
XX

```

```

PR 08-APR-1992; 92US-00866033.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Rota JS, Bellini WJ;
XX
DR WPI; 1993-351735/44.
DR N-PSDB; AAQ51095.
XX
XX Haemagglutinin and fusion glycoprotein of several wild-type measles
PT strains - used to construct vaccines for measles infection.
XX
PS Disclosure; Page 58-61; 119pp; English.
XX
XX HA (AAQ51088-94) and fusion glycoprotein (AAQ51095-97) sequences of
CC several wild-type measles strains are given. Shared amino acid variations
CC in wild-type measles glycoproteins are identified in five wild-type
CC measles viruses. A consensus polypeptide, the amino acid sequence of
CC which reflects variation common to more than one wild-type strain, is
CC the basis for constructing live attenuated vaccines, or recombinant
CC vaccines to replace older, less efficacious vaccines. Immunological
CC reagents useful in differentiating wild-type measles strains from other
CC known strains can also be produced. (Updated on 25-MAR-2003 to correct PN
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 550 AA;

Query Match 75.0%; Score 66; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.012;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 2 PDKLLTFIASDTCPV 17
Db 408 PDKILTYIAADHCPV 423
|||||:|:|:|:|:|
|||||:|:|:|:|:|

Search completed: September 28, 2005, 17:50:25
Job time : 92.25 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:56:03 ; Search time 87 Seconds  
(without alignments)  
79,534 Million cell updates/sec

Title: US-10-705-819B-6  
Perfect score: 83  
Sequence: 1 SHQYLVIKLIPNASLIE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1826554 seqs, 407025358 residues

Total number of hits satisfying chosen parameters: 1826554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	83	100.0	17	US-10-935-155A-6
2	83	100.0	17	Sequence 6, Appli
3	83	100.0	17	Sequence 6, Appli
4	83	100.0	662	US-10-705-819B-6
5	83	100.0	662	US-10-935-155A-27
6	76	91.6	662	US-10-705-819B-27
7	56	67.5	550	US-09-951-061A-141
8	52	62.7	15	US-10-161-097-40
9	51	61.4	550	US-09-873-233A-18
10	49	59.0	17	US-10-935-155A-6
11	49	59.0	17	US-10-705-819B-8

12	47	56.6	17	US-10-935-155A-9
13	47	56.6	17	US-10-705-819B-9
14	44	53.0	178	US-10-450-763-40992
15	44	53.0	256	US-09-988-837A-32
16	44	53.0	705	US-10-369-493-1864
17	43	51.8	53	US-10-424-599-170050
18	43	51.8	313	US-10-424-599-221101
19	43	51.8	336	US-10-437-963-156864
20	43	51.8	340	US-10-732-923-13971
21	42	50.6	40	US-10-424-599-244710
22	41	49.4	418	US-10-425-115-308956
23	41	49.4	2214	US-09-919-039-40
24	41	49.4	2214	US-10-176-847-94
25	41	49.4	2214	US-10-097-340-300
26	41	49.4	2214	US-10-464-368-89
27	41	49.4	2214	US-10-188-832-78
28	41	49.4	2214	US-10-473-127-810
29	41	49.4	2214	US-10-473-127-811
30	41	49.4	2214	US-10-473-127-812
31	41	49.4	2214	US-10-473-127-813
32	41	49.4	2214	US-10-473-127-814
33	41	49.4	2214	US-10-756-149-4693
34	41	49.4	2215	US-10-281-478-4
35	40	48.2	67	US-10-424-599-186040
36	40	48.2	80	US-10-425-115-189901
37	40	48.2	88	US-10-101-464A-758
38	40	48.2	88	US-10-864-252-758
39	40	48.2	321	US-10-433-757-10
40	40	48.2	571	US-09-071-035-168
41	40	48.2	571	US-10-206-576-168
42	40	48.2	571	US-10-912-362-168
43	40	48.2	593	US-09-071-035-166
44	40	48.2	593	US-10-206-576-166
45	40	48.2	593	US-10-912-362-166

ALIGNMENTS

RESULT 1

US-10-935-155A-6

; Sequence 6, Application US/10935155A

; Publication No. US20050037014A1

; GENERAL INFORMATION:

; APPLICANT: Jackson, David C

; APPLICANT: Ghosh, Souravi

; APPLICANT: Walker, John

; TITLE OF INVENTION: T Helper Cell Epitopes

; FILE REFERENCE: 47-212

; CURRENT APPLICATION NUMBER: US/10/935,155A

; CURRENT FILING DATE: 2004-09-08

; PRIOR APPLICATION NUMBER: US 09/890650

; PRIOR FILING DATE: 2002-03-22

; PRIOR APPLICATION NUMBER: US 10/705,819

; PRIOR FILING DATE: 2003-11-13

; PRIOR APPLICATION NUMBER: PCT/AU00/00070

; PRIOR FILING DATE: 2000-02-07

; PRIOR APPLICATION NUMBER: AU PP8533

; PRIOR FILING DATE: 1999-02-05

; PRIOR APPLICATION NUMBER: AU PQ2013

; PRIOR FILING DATE: 1999-08-04

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 17

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: PEPTIDE

; LOCATION: (1)-(17)

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: peptide P24

US-10-935-155A-6

Query Match 100.0%; Score 83; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-07; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||  
DB 1 SHQYLVKILPNASLIE 17

## RESULT 2

US-10-705-819B-6  
; Sequence 6, Application US/10705819B  
; Publication No. US20040058316A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-201  
; CURRENT APPLICATION NUMBER: US/10/705,819B  
; PRIOR FILING DATE: 2003-11-13  
; CURRENT FILING DATE: 2003-11-13  
; CURRENT APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P24  
US-10-705-819B-6

Query Match 100.0%; Score 83; DB 18; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1e-07; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||  
DB 1 SHQYLVKILPNASLIE 17

## RESULT 3

US-10-670-695-36  
; Sequence 36, Application US/10670695  
; Publication No. US20040058316A1  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Wayne A.  
; APPLICANT: Lappin, Michael R.  
; APPLICANT: Rosen, David K.  
; APPLICANT: Andrews, Janet S.  
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
; FILE REFERENCE: DI-9-1  
; CURRENT APPLICATION NUMBER: US/10/670,695  
; PRIOR FILING DATE: 2003-09-25  
; CURRENT APPLICATION NUMBER: 09/521,738  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 36  
; LENGTH: 662  
; TYPE: PRT

; ORGANISM: canine distemper virus  
US-10-670-695-36

Query Match 100.0%; Score 83; DB 15; Length 662;  
Best Local Similarity 100.0%; Pred. No. 6.6e-06; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||  
DB 162 SHQYLVKILPNASLIE 178

## RESULT 4

US-10-935-155A-27  
; Sequence 27, Application US/10935155A  
; Publication No. US20050037014A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-212  
; CURRENT APPLICATION NUMBER: US/10/935,155A  
; PRIOR FILING DATE: 2004-09-08  
; CURRENT APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 10/705,819  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(662)  
; OTHER INFORMATION: Description of Artificial Sequence: full length canine distemper  
; OTHER INFORMATION: fusion protein  
US-10-935-155A-27

Query Match 100.0%; Score 83; DB 17; Length 662;  
Best Local Similarity 100.0%; Pred. No. 6.6e-06; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Mismatches 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||  
DB 162 SHQYLVKILPNASLIE 178

## RESULT 5

US-10-705-819B-27  
; Sequence 27, Application US/10705819B  
; Publication No. US20050186220A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-201  
; CURRENT APPLICATION NUMBER: US/10/705,819B  
; PRIOR FILING DATE: 2003-11-13  
; CURRENT APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533



```
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(662)
; OTHER INFORMATION: Description of Artificial Sequence: full length canine distemper
; OTHER INFORMATION: fusion protein
US-10-705-819B-27

Query Match          100.0%; Score 83; DB 18; Length 662;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SHOYLVKILPNASLI 17
Db      162 SHOYLVKILPNASLI 178

RESULT 6
US-09-951-061A-141
; Sequence 141, Application US/09951061A
; Publication No. US20030082204A1
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Tartaglia, James
; APPLICANT: Taylor, Jill
; APPLICANT: Gettig, Russell
; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE
; TITLE OF INVENTION: RECOMBINANTS
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell, Boehnen, Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/951.061A
; FILING DATE: 13-SEP-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/354,138
; FILING DATE: 15-JUL-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/224,657
; FILING DATE: 16-APR-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,962
; FILING DATE: 08-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,867
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,283
; FILING DATE: 31-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
```

```
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/713,967
; FILING DATE: 11-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07,666,056
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-951-061A-141

Query Match          91.8%; Score 76; DB 10; Length 662;
Best Local Similarity 88.2%; Pred. No. 0.00012;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 SHOYLVKILPNASLI 17
Db      162 SHOYLVKILPNVSLIE 178

RESULT 7
US-09-873-233A-20
; Sequence 20, Application US/09873233A
; Patent No. US20020146434A1
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: GENE CODING FOR THE MEASLES VIRUS MUTANT ANTIGEN
; FILE REFERENCE: 0216-0451P
; CURRENT APPLICATION NUMBER: US/09/873,233A
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(550)
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-873-233A-20

Query Match          67.5%; Score 56; DB 9; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.36;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SHOYLVKILPNASLI 16
Db      50 SHOYLVKILPNITLL 65

RESULT 8
```

US-10-161-097-40  
; Sequence 40, Application US/10161097  
; Publication No. US20030096404A1  
; GENERAL INFORMATION:  
; APPLICANT: ROSENZWEIG, Michael  
; APPLICANT: PYKETT, Mark J.  
; APPLICANT: SCADDEN, David T.  
; APPLICANT: POZNANSKY, Mark C.  
; TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION  
; TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL  
; FILE REFERENCE: C1005/7012/KA/ERG  
; CURRENT APPLICATION NUMBER: US/10/161,097  
; CURRENT FILING DATE: 2002-05-31  
; PRIOR APPLICATION NUMBER: US/09/574,749  
; PRIOR FILING DATE: 2002-05-31  
; PRIOR APPLICATION NUMBER: US 60/107,972  
; PRIOR FILING DATE: 1998-11-12  
; PRIOR APPLICATION NUMBER: PCT/US99/26795  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: US 09/524,749  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 40  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Measles source  
US-10-161-097-40

Query Match 62.7%; Score 52; DB 14; Length 15;  
Best Local Similarity 66.7%; Pred. No. 0.032;  
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 HOYLVKLIIPNASLI 16  
DB 1 HQALVVKLIPNITLL 15  
||| ||||| :| :|

RESULT 9  
US-09-873-233A-18  
; Sequence 18, Application US/09873233A  
; Patent No. US20020146434A1  
; GENERAL INFORMATION:  
; APPLICANT: UEDA, Shigeharu  
; APPLICANT: WATANABE, Michiko  
; APPLICANT: KAWANISHI, Hitomi  
; TITLE OF INVENTION: GENE CODING FOR THE MEASLES VIRUS MUTANT ANTIGEN  
; FILE REFERENCE: 0216-0451P  
; CURRENT APPLICATION NUMBER: US/09/873,233A  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 550  
; TYPE: PRT  
; ORGANISM: Measles virus  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (1)...(550)  
; OTHER INFORMATION: any n or Xaa = Unknown  
US-09-873-233A-18

Query Match 61.4%; Score 51; DB 9; Length 550;  
Best Local Similarity 62.5%; Pred. No. 2.9;  
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHOYLVKLIIPNASLI 16  
DB 50 SHHSLVVKLIPNITLL 65  
||| ||||| :| :|

RESULT 10  
US-10-935-155A-8  
; Sequence 8, Application US/10935155A  
; Publication No. US20050037014A1  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, David C  
; APPLICANT: GHOSH, Souravi  
; APPLICANT: WALKER, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-212  
; CURRENT APPLICATION NUMBER: US/10/935,155A  
; CURRENT FILING DATE: 2004-09-08  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 10/705,819  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)...(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P23  
US-10-935-155A-8

Query Match 59.0%; Score 49; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOYLVKLI 10  
DB 8 SHOYLVKLI 17  
||| |||||

RESULT 11  
US-10-705-819B-8  
; Sequence 8, Application US/10705819B  
; Publication No. US20050186220A1  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, David C  
; APPLICANT: GHOSH, Souravi  
; APPLICANT: WALKER, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-201  
; CURRENT APPLICATION NUMBER: US/10/705,819B  
; CURRENT FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE

```
; LOCATION: (1)..(17)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v
; OTHER INFORMATION: Peptide P23
US-10-705-819B-8

Query Match          59.0%; Score 49; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLI 10
Db 8 SHQYLVIKLI 17

RESULT 12
US-10-935-155A-9
; Sequence 9, Application US/10935155A
; Publication No. US20050037014A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-212
; CURRENT APPLICATION NUMBER: US/10/935,155A
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 09/890650
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/705,819
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00070
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: AU P8533
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v
; OTHER INFORMATION: peptide P25
US-10-935-155A-9

Query Match          56.6%; Score 47; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KLIPNASLIE 17
Db 1 KLIPNASLIE 10

RESULT 13
US-10-705-819B-9
; Sequence 9, Application US/10705819B
; Publication No. US20050186220A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-201
; CURRENT APPLICATION NUMBER: US/10/705,819B
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 09/890650
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: PCT/AU00/00070
```

```
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: AU P8533
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v
; OTHER INFORMATION: peptide P25
US-10-705-819B-9

Query Match          56.6%; Score 47; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KLIPNASLIE 17
Db 1 KLIPNASLIE 10

RESULT 14
US-10-450-763-40992
; Sequence 40992, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40992
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (50)..(63)
; OTHER INFORMATION: CHOLINESTERASE SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00878F, p-value=9.723e-10, raw score of 5.37
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (2)..(163)
; OTHER INFORMATION: Carboxylesterases domain identified by Pfam, accession name
; OTHER INFORMATION: Coesterase, E-value=1.1e-15, Pfam score of 56.1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(178)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-40992

Query Match          53.0%; Score 44; DB 18; Length 178;
Best Local Similarity 43.8%; Pred. No. 14;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16
Db 95 AHQFLICYALPNESLL 110
```

RESULT 15  
US-09-898-837A-32  
; Sequence 32, Application US/09898837A  
; Publication NO. US20030077697A1  
; GENERAL INFORMATION:  
; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Vernet, Corine  
; APPLICANT: Herrmann, John L.  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Rastelli, Luca  
; APPLICANT: CuraGen Corporation  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Macdougall, John R  
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND  
; FILE REFERENCE: 15966-598 CIP  
; CURRENT APPLICATION NUMBER: US/09/898,837A  
; CURRENT FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-898-837A-32

Query Match 53.0%; Score 44; DB 10; Length 256;  
Best Local Similarity 41.2%; Pred. NO. 22;  
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLIE 17  
|:|:::| | | |:  
Db 71 SNVYILLEICPNGLME 87

Search completed: September 28, 2005, 18:26:03  
Job time : 87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:41:21 ; Search time 23.5 Seconds  
(without alignments)  
54.001 Million cell updates/sec

Title: US-10-705-819B-6  
Perfect score: 83  
Sequence: 1 SHQYLVKLPNASLIE 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	17	4	US-09-890-650-6
2	83	100.0	662	4	US-09-890-650-27
3	76	91.6	662	1	US-08-224-657-88
4	76	91.6	662	3	US-09-354-138-88
5	56	67.5	550	1	US-08-279-700-16
6	56	67.5	550	1	US-08-279-700-18
7	56	67.5	550	1	US-08-279-700-20
8	56	67.5	550	1	US-08-279-700-22
9	56	67.5	550	1	US-08-348-891A-5
10	56	67.5	550	2	US-08-905-817-5
11	56	67.5	550	3	US-09-230-944-20
12	56	67.5	550	4	US-09-873-233A-20
13	52	62.7	15	4	US-09-574-749B-40
14	51	61.4	550	3	US-09-230-944-18
15	51	61.4	550	4	US-08-873-233A-18
16	49	59.0	17	4	US-09-890-650-8
17	47	56.6	17	4	US-09-890-650-9
18	44	53.0	275	1	US-08-252-995D-13
19	44	53.0	275	2	US-08-834-108-13
20	41	49.4	95	4	US-09-248-796A-21274
21	41	49.4	2213	1	US-08-727-034-3
22	41	49.4	2214	1	US-08-727-034-7
23	41	49.4	2214	4	US-09-919-039-40
24	40	48.2	88	4	US-10-101-464A-758
25	40	48.2	330	4	US-09-248-796A-18497
26	40	48.2	387	4	US-09-248-796A-14674
27	40	48.2	562	4	US-09-134-000C-4308

28 40 48.2 571 4 US-09-071-035-168 Sequence 168, App  
29 40 48.2 575 4 US-09-252-991A-16697 Sequence 16697, A  
30 40 48.2 593 4 US-09-071-035-166 Sequence 166, App  
31 40 48.2 793 4 US-09-949-016-7810 Sequence 7810, App  
32 39 47.0 160 4 US-09-248-796A-28148 Sequence 28148, A  
33 39 47.0 260 4 US-09-248-796A-15296 Sequence 15296, A  
34 39 47.0 282 4 US-09-134-000C-6521 Sequence 6521, App  
35 39 47.0 284 4 US-09-270-767-41459 Sequence 41459, A  
36 39 47.0 281 4 US-10-101-464A-518 Sequence 518, App  
37 39 47.0 307 4 US-09-107-532A-4626 Sequence 4626, App  
38 39 47.0 308 4 US-09-107-532A-5653 Sequence 5653, App  
39 39 47.0 393 4 US-09-107-532A-4627 Sequence 4627, App  
40 39 47.0 441 4 US-09-107-532A-4623 Sequence 4623, App  
41 39 47.0 441 4 US-09-107-532A-4625 Sequence 4625, App  
42 39 47.0 685 4 US-10-101-464A-918 Sequence 918, App  
43 39 47.0 2291 4 US-09-822-871-2 Sequence 2, Appli  
44 38 45.8 208 4 US-09-107-532A-6611 Sequence 6611, App  
45 38 45.8 431 4 US-09-710-279-1474 Sequence 1474, App

## ALIGNMENTS

RESULT 1  
US-09-890-650-6  
; Sequence 6, Application US/09890650  
; Patent No. 6685947  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID CHARLES  
; APPLICANT: SOURAVI, GHOSH  
; APPLICANT: WALKER, JOHN  
; TITLE OF INVENTION: T HELPER CELL EPTOPES  
; FILE REFERENCE: 47-152  
; CURRENT APPLICATION NUMBER: US/09/890, 650  
; CURRENT FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: canine distemper virus  
US-09-890-650-6

Query Match 100.0%; Score 83; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.2e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHQYLVKLPNASLIE 17  
| | | | | | | | | | | | | | | | | | | |  
Db 1 SHQYLVKLPNASLIE 17

RESULT 2  
US-09-890-650-27  
; Sequence 27, Application US/09890650  
; Patent No. 6685947  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID CHARLES  
; APPLICANT: SOURAVI, GHOSH  
; APPLICANT: WALKER, JOHN  
; TITLE OF INVENTION: T HELPER CELL EPTOPES  
; FILE REFERENCE: 47-152  
; CURRENT APPLICATION NUMBER: US/09/890, 650  
; CURRENT FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: canine distemper virus  
US-09-890-650-27

Query Match 100.0%; Score 83; DB 4; Length 662;

Best Local Similarity 100.0%; Pred. No. 9e-07; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0; Indels 0; Gaps 0;

Qy 1 SHQYLVKLIIPNASLIE 17  
Db 162 SHQYLVKLIIPNASLIE 178

RESULT 3  
US-08-224-657-88  
; Sequence 88, Application US/08224657  
; Patent No. 5756102  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: Tartaglia, James  
; APPLICANT: Taylor, Jill  
; TITLE OF INVENTION: FOXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
; TITLE OF INVENTION: RECOMBINANTS  
; NUMBER OF SEQUENCES: 122  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/224,657  
; FILING DATE: 06-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2550  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; TELEX: 425066 CURTMS  
; INFORMATION FOR SEQ ID NO: 88:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 662 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-224-657-88

Query Match 91.6%; Score 76; DB 1; Length 662;  
Best Local Similarity 88.2%; Pred. No. 1.7e-05;  
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQYLVKLIIPNASLIE 17  
Db 162 SHQYLVKLIIPNASLIE 178

RESULT 4  
US-09-354-138-88  
; Sequence 88, Application US/09354138  
; Patent No. 6309647  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: Tartaglia, James  
; APPLICANT: Taylor, Jill  
; APPLICANT: Gettig, Russell  
; TITLE OF INVENTION: FOXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
; TITLE OF INVENTION: RECOMBINANTS  
; NUMBER OF SEQUENCES: 139  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue, 25th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/354,138  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/472,379  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US 08/416,646  
; FILING DATE: 05-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/224,657  
; FILING DATE: 16-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/073,962  
; FILING DATE: 08-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/776,867  
; FILING DATE: 23-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/621,614  
; FILING DATE: 30-NOV-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/938,283  
; FILING DATE: 31-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/105,483  
; FILING DATE: 12-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847,951  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/713,967  
; FILING DATE: 11-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07,666,056  
; FILING DATE: 07-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frommer, William S.  
; REGISTRATION NUMBER: 25,506  
; REFERENCE/DOCKET NUMBER: 454310-2860  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 840-3333  
; TELEFAX: (212) 840-0712  
; INFORMATION FOR SEQ ID NO: 88:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 662 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-09-354-138-88

Query Match 91.6%; Score 76; DB 3; Length 662;  
Best Local Similarity 88.2%; Pred. No. 1.7e-05;  
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQYLVKLIIPNASLIE 17

```
Db      162 SHQYLVIKLMNVSLIE 178
|||||:|||||
RESULT 5
US-08-279-700-16
; Sequence 16, Application US/08279700
; Patent No. 5578448
; GENERAL INFORMATION:
; APPLICANT: ROTA, Jennifer S.
; APPLICANT: BELLINI, William J.
; TITLE OF INVENTION: WILD-TYPE MEASLES VIRUS GLYCOPROTEINS:
; TITLE OF INVENTION: VACCINE AND DETECTION METHOD THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,700
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/866,033A
; FILING DATE: 19920408
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/140 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-279-700-16
Query Match      67.5%; Score 56; DB 1; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SHQYLVIKLI PNASLI 16
|||||:|||||
Db      50 SHQSLVIKLMNPNTLL 65
|||||:|||||

RESULT 7
US-08-279-700-20
; Sequence 20, Application US/08279700
; Patent No. 5578448
; GENERAL INFORMATION:
; APPLICANT: ROTA, Jennifer S.
; APPLICANT: BELLINI, William J.
; TITLE OF INVENTION: WILD-TYPE MEASLES VIRUS GLYCOPROTEINS:
; TITLE OF INVENTION: VACCINE AND DETECTION METHOD THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,700
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/866,033A
; FILING DATE: 19920408
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/140 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-279-700-16
Query Match      67.5%; Score 56; DB 1; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy      1 SHQYLVIKLI PNASLI 16
|||||:|||||
Db      50 SHQSLVIKLMNPNTLL 65
|||||:|||||

RESULT 6
US-08-279-700-18
; Sequence 18, Application US/08279700
; Patent No. 5578448
; GENERAL INFORMATION:
; APPLICANT: ROTA, Jennifer S.
; APPLICANT: BELLINI, William J.
; TITLE OF INVENTION: WILD-TYPE MEASLES VIRUS GLYCOPROTEINS:
; TITLE OF INVENTION: VACCINE AND DETECTION METHOD THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
```





```
; APPLICANT: MAKINO, Satoshi
; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE
; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
; TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,817
; FILING DATE: 04-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,891
; FILING DATE: 25-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,400
; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-293625
; FILING DATE: 14-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: KP-7501A
; TELEPHONE: 703-521-2297
; TELEFAX: 703-685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-905-817-5

Query Match 67.5%; Score 56; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16
Db 50 SHQSLVTKLMPNITLL 65

RESULT 11
US-09-230-944-20
; Sequence 20, Application US/09230944
; Patent No. 6277380
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: Measles Virus Mutant Antigen and Gene Coding for the
; TITLE OF INVENTION: Same
; FILE REFERENCE: 0216-0407P
; CURRENT APPLICATION NUMBER: US/09/230,944
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: PCT/JP98/02481
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 9-184285 JAPAN
; EARLIER FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 32

; APPLICANT: MAKINO, Satoshi
; TITLE OF INVENTION: ATTENUATED MEASLES VIRUS VACCINE
; TITLE OF INVENTION: CONTAINING SPECIFIC NUCLEOTIDE SEQUENCE AND A METHOD FOR
; TITLE OF INVENTION: ITS ABSOLUTE IDENTIFICATION
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: YOUNG & THOMPSON
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,817
; FILING DATE: 04-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,891
; FILING DATE: 25-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,400
; FILING DATE: 10-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 3-293625
; FILING DATE: 14-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: 32,925
; REFERENCE/DOCKET NUMBER: KP-7501A
; TELEPHONE: 703-521-2297
; TELEFAX: 703-685-0573
; TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-905-817-5

Query Match 67.5%; Score 56; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16
Db 50 SHQSLVTKLMPNITLL 65

RESULT 11
US-09-230-944-20
; Sequence 20, Application US/09230944
; Patent No. 6277380
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: Measles Virus Mutant Antigen and Gene Coding for the
; TITLE OF INVENTION: Same
; FILE REFERENCE: 0216-0407P
; CURRENT APPLICATION NUMBER: US/09/230,944
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: PCT/JP98/02481
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 9-184285 JAPAN
; EARLIER FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 32
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-230-944-20

Query Match 67.5%; Score 56; DB 3; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16
Db 50 SHQSLVTKLMPNITLL 65

RESULT 12
US-09-873-233A-20
; Sequence 20, Application US/09873233A
; Patent No. 6605284
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: GENE CODING FOR THE MEASLES VIRUS MUTANT ANTIGEN
; FILE REFERENCE: 0216-0451P
; CURRENT APPLICATION NUMBER: US/09/873,233A
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(550)
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-873-233A-20

Query Match 67.5%; Score 56; DB 4; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.066;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16
Db 50 SHQSLVTKLMPNITLL 65

RESULT 13
US-09-574-749B-40
; Sequence 40, Application US/09574749B
; Patent No. 6548299
; GENERAL INFORMATION:
; APPLICANT: ROSENZWEIG, Michael
; APPLICANT: PYKETT, Mark J.
; APPLICANT: SCADDEN, David T.
; APPLICANT: POZNANSKY, Mark C.
; TITLE OF INVENTION: LYMPHOID TISSUE-SPECIFIC CELL PRODUCTION
; TITLE OF INVENTION: FROM HEMATOPOIETIC PROGENITOR CELLS IN THREE-DIMENSIONAL
; FILE REFERENCE: C1005/7012/KA/ERG
; CURRENT APPLICATION NUMBER: US/09/574,749B
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/107,972
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: PCT/US99/26795
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: US 09/524,749
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 58
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Measles source
US-09-574-749B-40

Query Match          62.7%; Score 52; DB 4; Length 15;
Best Local Similarity 66.7%; Pred. No. 0.005;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 HQYLVIKLIPNASLI 16
Db 1 HQALVIKLPNITLL 15

RESULT 14
US-09-230-944-18
; Sequence 18, Application US/09230944
; Patent No. 6277380
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: Measles Virus Mutant Antigen and Gene Coding for the
; FILE REFERENCE: 0216-0407P
; CURRENT APPLICATION NUMBER: US/09/230,944
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: PCT/JP98/02481
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 9-184285 JAPAN
; EARLIER FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-230-944-18

Query Match          61.4%; Score 51; DB 3; Length 550;
Best Local Similarity 62.5%; Pred. No. 0.55;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVIKLIPNASLI 16
Db 50 SHSLVIKLPNITLL 65

RESULT 15
US-09-873-233A-18
; Sequence 18, Application US/09873233A
; Patent No. 6605284
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: GENE CODING FOR THE MEASLES VIRUS MUTANT ANTIGEN
; FILE REFERENCE: 0216-0451P
; CURRENT APPLICATION NUMBER: US/09/873,233A
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
```

```
; NAME/KEY: UNSURE
; LOCATION: (1)-(550)
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-873-233A-18

Query Match          61.4%; Score 51; DB 4; Length 550;
Best Local Similarity 62.5%; Pred. No. 0.55;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVIKLIPNASLI 16
Db 50 SHSLVIKLPNITLL 65

Search completed: September 28, 2005, 17:58:45
Job time : 24.5 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:40:37 ; Search time 17 Seconds  
(without alignments)  
96.217 Million cell updates/sec

Title: US-10-705-819B-6  
Perfect score: 83  
Sequence: 1 SHQYLVKILIPNASLIE 17  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*  
1: Pirl.\*  
2: Pirl2.\*  
3: Pirl3.\*  
4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	662	1 VGNZCD	cell fusion glycop
2	76	91.6	662	2 S21382	cell fusion protei
3	73	88.0	542	2 JQ2223	cell fusion protei
4	69	83.1	631	1 A48346	cell fusion glycop
5	69	83.1	631	1 VGNZPD	cell fusion glycop
6	68	81.9	552	2 S47034	cell fusion protei
7	57	68.7	534	1 JU0274	cell fusion glycop
8	56	67.5	550	1 R48556	cell fusion glycop
9	56	67.5	553	1 VGNZMV	cell fusion glycop
10	56	67.5	636	2 S47299	gene F protein - r
11	55	66.3	546	1 VGNZRK	cell fusion glycop
12	55	66.3	546	1 VGNZRL	cell fusion glycop
13	55	66.3	546	2 S47305	gene F protein - r
14	55	66.3	546	2 S47300	gene F protein - r
15	55	66.3	546	2 S55386	cell fusion protei
16	45	54.2	305	2 D69024	hypothetical prote
17	45	54.2	538	2 S52472	cell fusion protei
18	44	53.0	705	2 A48144	protein kinase CDC
19	43	51.8	538	1 VGNZNM	cell fusion glycop
20	43	51.8	538	1 VGNZMU	cell fusion glycop
21	42	50.6	150	2 AG1585	hypothetical prote
22	42	50.6	150	2 AH1231	hypothetical prote
23	42	50.6	219	2 AE0284	probable membrane
24	42	50.6	340	2 F75193	asparaginase (SC 3
25	42	50.6	355	2 S41686	geranylgeranyltran
26	42	50.6	538	1 B60004	cell fusion glycop
27	42	50.6	538	1 VGNZMS	cell fusion glycop
28	42	50.6	692	2 E96841	hypothetical prote
29	41	49.4	513	2 AE1921	hypothetical prote

30	41	49.4	1130	2	TL17399	probable DEAH ATP-
31	41	49.4	2215	2	T00348	Lrll protein - mou
32	40	48.2	206	2	AF1837	hypothetical prote
33	40	48.2	278	2	D82402	hypothetical prote
34	40	48.2	379	1	S43265	ubiquinol-cytochro
35	40	48.2	412	2	S23759	transposase - Leuc
36	40	48.2	521	2	T27896	hypothetical prote
37	40	48.2	554	2	E83618	malonate decarboxy
38	40	48.2	1413	2	D88844	protein 2K792.1 [i
39	40	48.2	2894	2	C64474	hypothetical prote
40	39.5	47.6	399	2	G97727	proton/sodium-glut
41	39.5	47.6	399	2	D71728	proton/sodium-glut
42	39	47.0	379	1	S43270	ubiquinol-cytochro
43	39	47.0	462	2	D84858	hypothetical prote
44	39	47.0	670	2	T32221	hypothetical prote
45	39	47.0	862	2	E36786	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

VGNZCD

cell fusion glycoprotein precursor - canine distemper virus

N;Contains: fusion protein F1; fusion protein F2

C;Species: canine distemper virus

C;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004

C;Accession: J50321

R;Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.

Virus Res 8, 373-386, 1987

A;Title: The nucleotide sequence of the gene encoding the F protein of canine distemper

A;Reference number: J50321; MUID:88129050; PMID:3433924

A;Accession: J50321

A;Molecule type: mRNA

A;Residues: 1-662 <BAR>

A;Cross-references: UNIPROT:P12569; GB:M21849; NID:G323241; PIDN:AAA42878.1; PID:G323242

C;Genetics:

A;Gene: F

C;Superfamily: parainfluenza virus cell fusion protein

C;Keywords: glycoprotein; membrane fusion; transmembrane protein

F;1-135/Domain: signal sequence #status predicted <SIG>

F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>

F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>

F;606-629/Domain: transmembrane #status predicted <MEW>

F;62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 83; DB 1; Length 662;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHQYLVKILIPNASLIE 17

Db 162 SHQYLVKILIPNASLIE 178

##### RESULT 2

S21382

cell fusion protein - canine distemper virus

C;Species: canine distemper virus

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S21382

R;Wild, T.F.; Bernard, A.; Spohner, D.; Valleval, D.; Drillien, R.

submitted to the EMBL Data Library, April 1992

A;Description: Vaccination of mice against canine distemper virus induced encephalitis w

A;Reference number: S21382

A;Accession: S21382

A;Status: preliminary

A;Molecule type: Genomic RNA

A;Residues: 1-662 <WIL>

A;Cross-references: UNIPROT:P12569; EMBL:X65509; NID:G58853; PIDN:CAA46481.1; PID:G58854

C;Superfamily: parainfluenza virus cell fusion protein

Query Match 91.6%; Score 76; DB 2; Length 662;

Best Local Similarity 88.2%; Pred. No. 3-2e-05;  
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||||:|:|:|  
Db 162 SHQYLVKILMPNVSLLIE 178

## RESULT 3

cell fusion glycoprotein F0 precursor - phocine distemper virus  
N;Contains: F1 and F2 chains  
C;Species: phocine distemper virus  
C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C;Accession: JQ2223  
R;Visser, I.K.G.; van der Heijden, R.W.J.; van de Bildt, M.W.G.; Kenter, M.J.H.; Oervell  
J. Gen. Virol. 74, 1989-1994, 1993  
A;Title: Fusion protein gene nucleotide sequence similarities, shared antigenic sites and  
e virus entity.  
A;Reference number: JQ2223; MUID:93389459; PMID:8376973  
A;Accession: JQ2223  
A;Molecule type: mRNA  
A;Residues: 1-542 <VIS>  
A;Cross-references: UNIPROT:Q7LZY1; GB:L07075  
A;Note: the authors translated the codon ATC for residue 4 as Leu  
C;Comment: This fusion protein F0 is cleaved into F1 and F2 chains.  
C;Genetics:

A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-15/Domain: signal sequence #status predicted <SIG>  
F;16-542/Product: fusion protein #status predicted <MAT>  
F;16-93/Product: F2 chain #status predicted <FPC>  
F;105-542/Product: F1 chain #status predicted <FIC>  
F;105-135/Region: hydrophobic  
F;486-512/Domain: transmembrane #status predicted <TM3>  
F;21,53,59,397/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.0%; Score 73; DB 2; Length 542;  
Best Local Similarity 82.4%; Pred. No. 8.7e-05;  
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||||:|:|:|  
Db 42 SHQYLVKILMPNVSLLID 58

## RESULT 4

cell fusion glycoprotein precursor - phocine distemper virus (strain Ulster/88)  
N;Contains: fusion protein F1; fusion protein F2  
C;Species: phocine distemper virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: A48346  
R;Curran, M.D.; Lu, Y.J.; Rima, B.K.  
Arch. Virol. 126, 159-169, 1992  
A;Title: The fusion protein gene of phocine distemper virus: nucleotide and deduced amino  
acid sequence  
A;Reference number: A48346; MUID:92398437; PMID:1524494  
A;Accession: A48346  
A;Molecule type: mRNA  
A;Residues: 1-631 <CUR>  
A;Cross-references: UNIPROT:P28886  
A;Note: sequence extracted from NCBI backbone (NCBIN:113098, NCBI:P113099)  
C;Genetics:

A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>  
F;89-106/Domain: transmembrane #status predicted <TM1>  
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>  
F;194-219/Domain: transmembrane #status predicted <TM2>  
F;575-595/Domain: transmembrane #status predicted <TM3>  
F;110,142/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.1%; Score 69; DB 1; Length 631;  
Best Local Similarity 70.6%; Pred. No. 0.00051;  
Matches 12; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||||:|:|:|  
Db 131 SHQYLVKILMPNVSIIID 147

## RESULT 5

VGNZPD  
cell fusion glycoprotein precursor - phocine distemper virus  
N;Contains: fusion protein F1; fusion protein F2  
C;Species: phocine distemper virus  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
C;Accession: JQ1368  
R;Koewames, J.; Blixenkrone-Moeller, M.; Sharma, B.; Oervell, C.; Norrby, E.  
J. Gen. Virol. 72, 2959-2966, 1991  
A;Title: The nucleotide sequence and deduced amino acid composition of the haemagglutinin  
of the phocine distemper virus  
A;Reference number: JQ1368; MUID:92113538; PMID:1765768  
A;Accession: JQ1368  
A;Molecule type: genomic RNA  
A;Residues: 1-631 <KOV>  
A;Cross-references: UNIPROT:P28886  
C;Genetics:

A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-188/Product: cell fusion glycoprotein F2 #status predicted <FP2>  
F;89-106/Domain: transmembrane #status predicted <TM1>  
F;189-193/Region: cleavage processing #status predicted  
F;194-631/Product: cell fusion glycoprotein F1 #status predicted <FP1>  
F;194-212/Domain: transmembrane #status predicted <TM2>  
F;575-595/Domain: transmembrane #status predicted <TM3>  
F;110,142,148,486/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.1%; Score 69; DB 1; Length 631;  
Best Local Similarity 70.6%; Pred. No. 0.00051;  
Matches 12; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||||:|:|:|  
Db 131 SHQYLVKILMPNVSIIID 147

## RESULT 6

S47034  
cell fusion protein precursor - porpoise morbillivirus  
N;Alternate names: F protein  
C;Species: porpoise morbillivirus  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S47034  
R;Bolt, G.; Gottschalk, E.; Blixenkrone-Moeller, M.; Wishaupt, R.G.A.; Welsh, M.J.; Bar  
submitted to the EMBL Data Library, July 1994  
A;Description: Nucleotide sequence comparisons of the F and M genes of cetacean morbillivir  
uses  
A;Reference number: S47034  
A;Accession: S47034  
A;Molecule type: mRNA  
A;Residues: 1-552 <BOL>  
A;Cross-references: UNIPROT:Q66147; EMBL:X80757; NID:g520639; PIDN:CAA56731.1; PID:g52064  
A;Experimental source: isolate Ulster 88  
A;Note: the source is designated as Cetacean morbillivirus  
C;Superfamily: parainfluenza virus cell fusion protein  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;26-552/Product: fusion protein #status predicted <MAT>

Query Match 81.9%; Score 68; DB 2; Length 552;  
Best Local Similarity 70.6%; Pred. No. 0.00067;  
Matches 12; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
|||||||:|:|:|

Db 52 SHQYLVIKLPNVTMID 68

## RESULT 7

JU0274  
cell fusion glycoprotein precursor - subacute sclerosing panencephalitis virus (strain Y  
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C;Species: subacute sclerosing panencephalitis virus, SSPRV  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jun-2000  
C;Accession: JU0274  
R;Komase, K.; Haga, T.; Yoshikawa, Y.; Sato, T.A.; Yamanouchi, K.  
Virus Genes 4, 173-181, 1990  
A;Title: Molecular analysis of structural protein genes of the Yamagata-1 strain of defe  
A;Reference number: JU0274; MUID:90385702; PMID:11698327  
A;Accession: JU0274  
A;Molecule type: mRNA  
A;Residues: 1-534 <KOM>  
A;Cross-references: EMBL:DI0548; NID:q222256; PIDN:BA001405.1; PID:q222257  
A;Note: the authors translated the codon GTA for residue 459 as Gly and GGG for residue  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F;108-534/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F;498-514/Domain: transmembrane #status predicted <TMN>  
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.7%; Score 57; DB 1; Length 534;  
Best Local Similarity 68.8%; Pred. No. 0.055;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQYLVIKLPNASLI 16  
|||||:|:|:  
Db 50 SHQSLVILKLPNTITLL 65

## RESULT 8

E48556  
cell fusion glycoprotein precursor - measles virus (strain AIK-C)  
C;Species: measles virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: E48556  
R;Mori, T.; Sasaki, K.; Hashimoto, H.; Makino, S.  
Virus Genes 7, 67-81, 1993  
A;Title: Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C  
A;Reference number: E48556; MUID:93227570; PMID:8470368  
A;Accession: E48556  
A;Molecule type: genomic RNA  
A;Residues: 1-550 <MOR>  
A;Cross-references: UNIPROT:P35973; GB:S58435; NID:q299460; PIDN:AAB26145.1; PID:g299465  
A;Note: sequence extracted from NCBI backbone (NCBIN:129264, NCBI:P.129272)  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-107/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F;108-550/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F;113-138/Region: hydrophobic  
F;495-514/Domain: transmembrane #status predicted <TMN>  
F;6,29,61,67/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.5%; Score 56; DB 1; Length 550;  
Best Local Similarity 68.8%; Pred. No. 0.086;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQYLVIKLPNASLI 16  
|||||:|:|:  
Db 50 SHQSLVILKLPNTITLL 65

## RESULT 9

VGNZMV  
cell fusion glycoprotein precursor - measles virus  
C;Species: measles virus  
C;Date: 31-Mar-1988 #sequence\_revision 31-Mar-1989 #text\_change 09-Jul-2004  
C;Accession: A26962; A25616; PQ0384  
R;Buckland, R.; Gerald, C.; Barker, R.; Wild, T.F.  
J. Gen. Virol. 68, 1695-1703, 1987  
A;Title: Fusion glycoprotein of measles virus: nucleotide sequence of the gene and compa  
A;Reference number: A92794; MUID:87224816; PMID:3585281  
A;Accession: A26962  
A;Molecule type: mRNA  
A;Residues: 1-553 <BUC>  
A;Cross-references: UNIPROT:O93055; GB:D00090; NID:q222061; PIDN:BA00056.1; PID:g222062  
A;Experimental source: strain Halle  
R;Richardson, C.; Hull, D.; Greer, P.; Hasel, K.; Berkovich, A.; Englund, G.; Bellini, W  
Virology 155, 508-523, 1986  
A;Title: The nucleotide sequence of the mRNA encoding the fusion protein of measles virus  
A;Reference number: A94350; MUID:87071668; PMID:3788062  
A;Accession: A25616  
A;Molecule type: mRNA  
A;Residues: 4-553 <RIC>  
A;Cross-references: GB:M14915; NID:g331762; PIDN:AAA46423.1; PID:g331763  
A;Experimental source: strain Edmonston  
R;Schulz, T.F.; Hoad, J.G.; Whitby, D.; Tizard, E.J.; Dillon, M.J.; Weiss, R.A.  
J. Gen. Virol. 73, 1581-1586, 1992  
A;Title: A measles virus isolate from a child with Kawasaki disease: sequence comparison  
A;Reference number: PQ0374; MUID:92300360; PMID:1607874  
A;Accession: PQ0380  
A;Molecule type: genomic RNA  
A;Residues: 272-553 <SCH1>  
A;Experimental source: isolate CL  
A;Accession: PQ0384  
A;Molecule type: genomic RNA  
A;Residues: 272-553 <SCH2>  
A;Experimental source: isolate SE  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-25/Domain: signal sequence #status predicted <SIG>  
F;26-110/Product: cell fusion glycoprotein F2 #status predicted <FF2>  
F;111-553/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F;501-517/Domain: transmembrane #status predicted <TMN>  
F;32,64,70/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.5%; Score 56; DB 1; Length 553;  
Best Local Similarity 68.8%; Pred. No. 0.086;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQYLVIKLPNASLI 16  
|||||:|:|:  
Db 53 SHQSLVILKLPNTITLL 68

## RESULT 10

S47299  
gene F protein - rinderpest virus  
C;Species: rinderpest virus  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S47299  
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
submitted to the EMBL Data Library, March 1994  
A;Description: The complete nucleotide sequence of the fusion protein gene of the vaccini  
A;Reference number: S47299  
A;Accession: S47299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-636 <EVA>  
A;Cross-references: UNIPROT:Q86486; EMBL:Z31655; NID:q535391; PIDN:CAA83481.1; PID:g53535  
C;Superfamily: parainfluenza virus cell fusion protein

Query Match 67.5%; Score 56; DB 2; Length 636;

Best Local Similarity 64.7%; Pred. No. 0.13;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
||| ||||| : : :  
Db 136 SHQSLVILKLPNTTAID 152

## RESULT 11

VGNZRK  
cell fusion glycoprotein precursor - rinderpest virus (strain Kabete O)  
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C;Species: rinderpest virus  
C;Date: 31-Dec-1989 #sequence\_revision 11-Dec-1989 #text\_change 09-Jul-2004  
C;Accession: A31051  
R;Hsu, D.; Yamanaka, M.; Miller, J.; Dale, B.; Grubman, M.; Yilma, T.  
Virology 166, 149-153, 1988  
A;Title: Cloning of the fusion gene of rinderpest virus: comparative sequence analysis  
A;Reference number: A31051; MUID:88322864; PMID:3413983  
A;Accession: A31051  
A;Molecule type: genomic RNA  
A;Residues: 1-546 <HSU>  
A;Cross-references: UNIPROT:P12574  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-108/Product: cell fusion glycoprotein F2 #status predicted <FF1>  
F;109-546/Product: cell fusion glycoprotein F1 #status predicted <FF2>  
F;109-134/Domain: transmembrane #status predicted <TN1>  
F;491-513/Domain: transmembrane #status predicted <TN2>  
F;25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.3%; Score 55; DB 1; Length 546;  
Best Local Similarity 64.7%; Pred. No. 0.13;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
||| ||||| : : :  
Db 46 SHQTLVILKLPNTTAID 62

## RESULT 12

VGNZRL  
cell fusion glycoprotein precursor - rinderpest virus (strain L)  
N;Contains: fusion glycoprotein F1; fusion glycoprotein F2  
C;Species: rinderpest virus  
C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 09-Jul-2004  
C;Accession: A28921  
R;Tsukiyama, K.; Yoshikawa, Y.; Yamanouchi, K.  
Virology 164, 523-530, 1988  
A;Title: Fusion glycoprotein (F) of rinderpest virus: entire nucleotide sequence of the  
A;Reference number: A28921; MUID:88219541; PMID:3285575  
A;Accession: A28921  
A;Molecule type: mRNA  
A;Residues: 1-546 <TSU>  
A;Cross-references: UNIPROT:P10864; GB:M20870; NID:G333898; PIDN:AAA47399.1; PID:G333899  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-104/Product: cell fusion glycoprotein F2 #status predicted <F2>  
F;105-546/Product: cell fusion glycoprotein F1 #status predicted <FF1>  
F;109-133/Domain: transmembrane #status predicted <TN1>  
F;485-513/Domain: transmembrane #status predicted <TN2>  
F;25,57,63/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.3%; Score 55; DB 1; Length 546;  
Best Local Similarity 64.7%; Pred. No. 0.13;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
||| ||||| : : :  
Db 46 SHQSLVILKLPNTTAID 62

## RESULT 13

S47305  
gene F protein - rinderpest virus  
C;Species: rinderpest virus  
C;Date: 20-Oct-1994 #sequence\_revision 08-Sep-1995 #text\_change 09-Jul-2004  
C;Accession: S47305; S47301  
R;Baron, M.D.; Barrett, T.  
Submitted to the EMBL Data Library, March 1994  
A;Description: The sequence of the N and L genes of Rinderpest virus, and the 50 and 30  
A;Reference number: S47283  
A;Accession: S47305  
A;Molecule type: mRNA  
A;Residues: 1-546 <BAR>  
A;Cross-references: UNIPROT:P41356; EMBL:Z30697; NID:G535396; PIDN:CAA83181.1; PID:G5354  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: transmembrane protein

Query Match 66.3%; Score 55; DB 2; Length 546;  
Best Local Similarity 64.7%; Pred. No. 0.13;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
||| ||||| : : :  
Db 46 SHQTLVILKLPNTTAID 62

## RESULT 14

S47300  
gene F protein - rinderpest virus  
C;Species: rinderpest virus  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S47300; PQ0865  
R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.  
Submitted to the EMBL Data Library, March 1994  
A;Description: the complete nucleotide sequence of the fusion protein gene of the vaccine  
A;Reference number: S47299  
A;Accession: S47300  
A;Molecule type: DNA  
A;Residues: 1-546 <EVA>  
A;Cross-references: UNIPROT:P41360; EMBL:Z31656; NID:G535406; PIDN:CAA83482.1; PID:G5354  
R;Chamberlain, R.W.; Mamwayi, H.M.; Hockley, E.; Shalla, M.S.; Goatley, L.; Knowles, N.J.  
J. Gen. Virol. 74, 2775-2780, 1993  
A;Title: Evidence for different lineages of rinderpest virus reflecting their geographic  
A;Reference number: PQ0865; MUID:94103786; PMID:8277286  
A;Accession: PQ0865  
A;Molecule type: mRNA  
A;Residues: 86-191 <CHA>  
C;Genetics:  
A;Gene: F  
C;Superfamily: parainfluenza virus cell fusion protein  
C;Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 66.3%; Score 55; DB 2; Length 546;  
Best Local Similarity 64.7%; Pred. No. 0.13;  
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17  
||| ||||| : : :  
Db 46 SHQSLVILKLPNTTAID 62

## RESULT 15

S55386  
cell fusion protein - peste-des-petits-ruminants virus (strain 75/1)  
N;Alternate names: F protein  
C;Species: peste-des-petits-ruminants virus  
A;Variety: Strain 75/1  
C;Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 09-Jul-2004

Search completed: September 28, 2005, 17:57:06  
Job time : 18 secs

***This Page Blank (usoto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:39:51; Search time 80.5 Seconds  
(without alignments)  
108.141 Million cell updates/sec

Title: US-10-705-819B-6

Perfect score: 83

Sequence: 1 SHQYLVIKILIPNASLIE 17

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	100.0	662	1 VGLF CDVO	P12569 canine dist
2	76	91.6	111	2 O71064	O71064 canine dist
3	76	91.6	111	2 Q7TES1	Q7TES1 canine dist
4	76	91.6	111	2 Q7TH85	Q7TH85 canine dist
5	76	91.6	111	2 Q7TH86	Q7TH86 canine dist
6	76	91.6	662	2 O89327	O89327 canine dist
7	76	91.6	662	2 O6SYV0	O6SYV0 canine dist
8	76	91.6	662	2 Q6T1B5	Q6T1B5 canine dist
9	76	91.6	662	2 Q9DX22	Q9DX22 canine dist
10	73	88.0	81	2 Q9Q086	Q9Q086 canine dist
11	73	88.0	111	2 Q71057	Q71057 canine dist
12	73	88.0	111	2 Q71058	Q71058 canine dist
13	73	88.0	111	2 Q71061	Q71061 canine dist
14	73	88.0	111	2 Q71063	Q71063 canine dist
15	73	88.0	111	2 Q71065	Q71065 canine dist
16	73	88.0	111	2 Q92764	Q92764 canine dist
17	73	88.0	111	2 Q92765	Q92765 canine dist
18	73	88.0	111	2 Q7TES0	Q7TES0 canine dist
19	73	88.0	111	2 Q7TH83	Q7TH83 canine dist
20	73	88.0	111	2 Q7TH84	Q7TH84 canine dist
21	73	88.0	111	2 Q9WLY0	Q9WLY0 canine dist
22	73	88.0	111	2 Q9WLY1	Q9WLY1 canine dist
23	73	88.0	111	2 Q9WLY2	Q9WLY2 canine dist
24	73	88.0	266	2 O55333	O55333 canine dist
25	73	88.0	266	2 O55334	O55334 canine dist
26	73	88.0	266	2 O55335	O55335 canine dist
27	73	88.0	266	2 O55336	O55336 canine dist
28	73	88.0	266	2 O55337	O55337 canine dist
29	73	88.0	266	2 O55338	O55338 canine dist
30	73	88.0	530	2 Q8QV06	Q8QV06 canine dist
31	73	88.0	542	2 Q7LZT1	Q7LZT1 phocine dis

32	73	88.0	662	2	O91KN3	O91KN3 canine dist
33	73	88.0	662	2	Q68PH9	Q68PH9 canine dist
34	73	88.0	662	2	Q6TV24	Q6TV24 canine dist
35	73	88.0	662	2	Q6TV32	Q6TV32 canine dist
36	73	88.0	662	2	Q9YKL7	Q9YKL7 canine dist
37	71	85.5	111	2	O9WNU7	O9WNU7 canine dist
38	70	84.3	111	2	O71059	O71059 canine dist
39	70	84.3	111	2	O71060	O71060 canine dist
40	70	84.3	111	2	O71062	O71062 canine dist
41	69	83.1	631	1	VGLF_PHODV	P28886 phocine dis
42	68	81.9	552	2	O56852	O56852 dolphin mor
43	68	81.9	552	2	O66147	O66147 cetacean mor
44	65	78.3	552	2	O66409	O66409 dolphin mor
45	59	71.1	553	2	O91248	O91248 measles vir

#### ALIGNMENTS

##### RESULT 1

ID	VGLF CDVO	STANDARD;	PRT;	662 AA.
AC	P12569; Q65991;			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;			
DE	Fusion glycoprotein F1]			
GN	Name=F;			
OS	Canine distemper virus (strain Onderstepoort) (CDV).			
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;			
OC	Paramyxoviridae; Paramyxovirinae; Morbilliviruses.			
OX	NCBI_TaxID=11233;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MDLINE=88129050; PubMed=3433924; DOI=10.1016/0168-1702(87)90009-8;			
RA	Barrett T., Clarke D.K., Evans S.A., Rima B.K.;			
RT	"the nucleotide sequence of the gene encoding the F protein of canine distemper virus: a comparison of the deduced amino acid sequence with other paramyxoviruses.";			
RL	Virus Res. 8:373-386(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MDLINE=9327696; PubMed=8470428; DOI=10.1016/0264-410X(93)90285-6;			
RA	Wild T.F., Bernard A., Spehner D., Villevall D., Drillion R.;			
RT	"Vaccination of mice against canine distemper virus-induced encephalitis with vaccinia virus recombinants encoding measles or canine distemper virus antigens.";			
RL	Vaccine 11:438-444(1993).			
CC	- - FUNCTION: This protein directs fusion of viral and cellular membranes.			
CC	- - SUBUNIT: Heterodimer of F1 and F2; disulfide-linked.			
CC	- - SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein family.			

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC	EMBL; M21849; AAA42878.1; -;	
DR	EMBL; X65503; CAA46481.1; -;	
DR	PIR; JS0321; VGNZCD.	
DR	PIR; S21382; S21382.	
DR	HSSP; P04849; 1SVF.	
DR	InterPro; IPR000776; Fusion gly.	
DR	Pfam; PF00523; Fusion gly; 1.	
KW	Envelope protein; Fusion protein; Glycoprotein; Signal; Transmembrane.	
FT	SIGNAL 1 ? 662 Fusion glycoprotein F0.	
FT	CHAIN	

```

FT CHAIN ? 224 Fusion glycoprotein F2.
FT CHAIN 225 662 Fusion glycoprotein F1.
FT TRANSMEM 606 629 Potential.
FT DISULFID 180 307 Linkage between F2 and F1 (Potential).
FT CARBOHYD 62 62 N-linked (GLCNAC. .) (Potential).
FT CARBOHYD 141 141 N-linked (GLCNAC. .) (Potential).
FT CARBOHYD 173 173 N-linked (GLCNAC. .) (Potential).
FT CARBOHYD 179 179 N-linked (GLCNAC. .) (Potential).
FT CONFLICT 3 3 R -> K (in Ref. 2).
FT CONFLICT 140 140 D -> N (in Ref. 2).
FT CONFLICT 152 152 I -> S (in Ref. 2).
FT CONFLICT 171 171 I -> M (in Ref. 2).
FT CONFLICT 174 174 A -> V (in Ref. 2).
FT CONFLICT 662 662 L -> H (in Ref. 2).
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 100.0%; Score 83; DB 1; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17
Db 162 SHQYLVKILPNASLIE 178
|||||

RESULT 2
ID 071064 PRELIMINARY; PRT; 111 AA.
AC 071064;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF026232; AAC08754.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
FT NON_TER 1 1
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 12065 MW; 1B2DA226E5131AP9 CRC64;

Query Match 91.6%; Score 76; DB 2; Length 111;
Best Local Similarity 88.2%; Pred. No. 3.2e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17
Db 7 SHQYLVKILPNVSLIE 23
|||||

RESULT 3
ID 07TES1 PRELIMINARY; PRT; 111 AA.
AC 07TES1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.

```

```

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=lederle;
RA Lednicky J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,
RA Sarich N.A., Witecki K.E., Braid M.D., Pedrak C., Houde C.M.;
RT "Effective primary isolation of wild-type Canine distemper virus in
MDCK, MVI lu and Vero cells without nucleotide sequence changes within
the entire haemagglutinin protein gene and in subgenomic sections of
the fusion and phosphoprotein genes.";
RL J. Virol. Methods 118:147-157(2004).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
CC EMBL; AY288311; AAP58259.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
FT NON_TER 1 1
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11976 MW; F69280F6ABB74BAF CRC64;

Query Match 91.6%; Score 76; DB 2; Length 111;
Best Local Similarity 88.2%; Pred. No. 3.2e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17
Db 7 SHQYLVKILPNVSLIE 23
|||||

RESULT 4
Q7TH85 PRELIMINARY; PRT; 111 AA.
ID Q7TH85
AC Q7TH85;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
SEQUENCE FROM N.A.
RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;
RA Lednicky J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,
RA Sarich N.A., Witecki K.E., Braid M.D., Pedrak C., Houde C.M.;
RT "Effective primary isolation of wild-type Canine distemper virus in
MDCK, MVI lu and Vero cells without nucleotide sequence changes within
the entire haemagglutinin protein gene and in subgenomic sections of
the fusion and phosphoprotein genes.";
RL J. Virol. Methods 118:147-157(2004).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
CC EMBL; AY288311; AAP44532.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
FT NON_TER 1 1
FT NON_TER 111 111
SQ SEQUENCE 111 AA; 11933 MW; 9E570321B1D4108 CRC64;

```

Query Match 91.6%; Score 76; DB 2; Length 111;  
 Best Local Similarity 88.2%; Pred. No. 3.2e-05;  
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHQYLVKILIPNASLIE 17  
 |||||:|||||  
 Db 7 SHQYLVKILIPNASLIE 23

RESULT 5  
 Q7TH86 PRELIMINARY; PRT; 111 AA.  
 AC Q7TH86;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Fusion protein (Fragment).  
 OS Canine distemper virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;  
 RA Lednický J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,  
 RA Sarich N.A., Wittecki K.E., Braid M.D., Pedrak C., Houde C.M.;  
 RT "Effective primary isolation of wild-type Canine distemper virus in  
 MDCK, Mv1 Lu and Vero cells without nucleotide sequence changes within  
 RT the entire haemagglutinin protein gene and in subgenomic sections of  
 RT the fusion and phosphoprotein genes.";   
 RL J. Virol. Methods 118:147-157(2004)  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 CC family.  
 DR EMBL: AY289612; AAP44531.1; -.  
 DR GO: GO:0019031; C:viral envelope; IEA.  
 DR GO: GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro: IPR000776; Fusion\_gly.  
 DR Pfam: PF00523; Fusion\_gly; 1.  
 DR Envelope protein; Fusion protein.  
 FT NON TER 1 111  
 FT NON TER 111 111  
 SQ SEQUENCE 111 AA; 11976 MW; F69280F6ABB74BAF CRC64;

Query Match 91.6%; Score 76; DB 2; Length 111;  
 Best Local Similarity 88.2%; Pred. No. 3.2e-05;  
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQYLVKILIPNASLIE 17  
 |||||:|||||  
 Db 7 SHQYLVKILIPNASLIE 23

RESULT 6  
 O89327 PRELIMINARY; PRT; 662 AA.  
 AC O89327;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Fusion protein F.  
 OS Canine distemper virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93174978; PubMed=8438593;  
 RA Sidhu M.S., Husar W., Cook S.D., Dowling P.C., Udem S.A.;  
 RT "Canine distemper terminal and intergenic non-protein coding  
 RT nucleotide sequences: completion of the entire CDV genome sequence.";   
 RL Virology 193:66-72(1993).

RN SEQUENCE FROM N.A.  
 RP Sidhu M.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 CC family.  
 DR EMBL: AF014953; AAC26994.1; -.  
 DR HSSP; P04849; 1SVF.  
 DR GO: GO:0019031; C:viral envelope; IEA.  
 DR GO: GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro: IPR000776; Fusion\_gly.  
 DR InterPro: IPR009050; Globin-like.  
 DR Pfam: PF00523; Fusion\_gly; 1.  
 KW Envelope protein; Fusion protein.  
 SQ SEQUENCE 662 AA; 72951 MW; 80B144C6B9801898 CRC64;

Query Match 91.6%; Score 76; DB 2; Length 662;  
 Best Local Similarity 88.2%; Pred. No. 0.00019;  
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQYLVKILIPNASLIE 17  
 |||||:|||||  
 Db 162 SHQYLVKILIPNASLIE 178

RESULT 7  
 Q6SYVO PRELIMINARY; PRT; 662 AA.  
 ID Q6SYVO  
 AC Q6SYVO;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Fusion protein F.  
 OS Canine distemper virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lednický J.A., Meehan T.P., Sarich N.A.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;  
 RA Lednický J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,  
 RA Sarich N.A., Wittecki K.E., Braid M.D., Pedrak C., Houde C.M.;  
 RT "Effective primary isolation of wild-type Canine distemper virus in  
 MDCK, Mv1 Lu and Vero cells without nucleotide sequence changes within  
 RT the entire haemagglutinin protein gene and in subgenomic sections of  
 RT the fusion and phosphoprotein genes.";   
 RL J. Virol. Methods 118:147-157(2004)  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 CC family.  
 DR EMBL: AY445077; AAR16539.1; -.  
 DR EMBL: AY542312; AAS48411.1; -.  
 DR EMBL: AY466011; AAR30102.1; -.  
 DR GO: GO:0019031; C:viral envelope; IEA.  
 DR GO: GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro: IPR000776; Fusion\_gly.  
 DR InterPro: IPR009050; Globin-like.  
 DR Pfam: PF00523; Fusion\_gly; 1.  
 KW Envelope protein; Fusion protein.  
 SQ SEQUENCE 662 AA; 72736 MW; 2DD1B45291D3D315 CRC64;

Query Match 91.6%; Score 76; DB 2; Length 662;  
 Best Local Similarity 88.2%; Pred. No. 0.00019;  
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SHQYLVKILIPNASLIE 17

```

|||||:|:|:|:|:|
Db 162 SHOYLVIKLPNVSLIE 178

RESULT 8
Q6TIB5
ID Q6TIB5 PRELIMINARY; PRT; 662 AA.
AC Q6TIB5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;
RA Lednický J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,
RA Sarich N.A., Witecki K.E., Braid M.D., Pedrak C., Houde C.M.;
RT "Effective primary isolation of wild-type Canine distemper virus in
RT MDCK, Mv1 Lu and Vero cells without nucleotide sequence changes within
RT the entire haemagglutinin protein gene and in subgenomic sections of
RT the fusion and phosphoprotein genes.";
RL J. Virol. Methods 118:147-157(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Lednický J.A., Meehan T.P., Sarich N.A.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY443350; AAR32272.1; -.
DR EMBL; AY395984; AAO96334.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72544 MW; E7BA3BD7D734B77 CRC64;

Query Match 91.6%; Score 76; DB 2; Length 662;
Best Local Similarity 88.2%; Pred. No. 0.00019;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOYLVIKLPNVSLIE 17
Db 162 SHOYLVIKLPNVSLIE 178

RESULT 10
Q9QQ86
ID Q9QQ86 PRELIMINARY; PRT; 81 AA.
AC Q9QQ86;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20293475; PubMed=10831852; DOI=10.1016/S0378-1135(00)00192-9;
RA Frollich K., Czupalla O., Haas L., Hentschke J., Dedek J., Fickel J.;
RT "Epizootiological investigations of canine distemper virus in free-
RT ranging carnivores from Germany.";
RL Vet. Microbiol. 74:283-292(2000).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF158979; AAD47252.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
DR Envelope protein; Fusion protein.
FT NON_TER 1
FT NON_TER 81
SQ SEQUENCE 81 AA; 8768 MW; 7791FF09B40DEB6B CRC64;

Query Match 88.0%; Score 73; DB 2; Length 81;
Best Local Similarity 82.4%; Pred. No. 7.9e-05;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOYLVIKLPNVSLIE 17
Db 3 SHOYLVIKLPNVSLID 19

RESULT 11
O71057
ID O71057 PRELIMINARY; PRT; 111 AA.
AC O71057;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

```

```

OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026220; AAC08742.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON TER 1 111
FT NON TER 1 111
SQ SEQUENCE 111 AA; 11937 MW; 985E0976685C544A CRC64;

Query Match 88.0%; Score 73; DB 2; Length 111;
Best Local Similarity 82.4%; Pred. No. 0.00011;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOYLVKIPNASLIE 17
DB 7 SHOYLVKIPNVSLID 23

RESULT 12
O71058
ID O71058 PRELIMINARY; PRT; 111 AA.
AC O71058;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026221; AAC08743.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON TER 1 111
FT NON TER 1 111
SQ SEQUENCE 111 AA; 11923 MW; 71405DAD7C24C8A8 CRC64;

Query Match 88.0%; Score 73; DB 2; Length 111;
Best Local Similarity 82.4%; Pred. No. 0.00011;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOYLVKIPNASLIE 17
DB 7 SHOYLVKIPNVSLID 23

RESULT 13
O71061
ID O71061 PRELIMINARY; PRT; 111 AA.
AC O71061;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)

```

```

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026227; AAC08749.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON TER 1 111
FT NON TER 1 111
SQ SEQUENCE 111 AA; 11879 MW; 71469C067AF913A8 CRC64;

Query Match 88.0%; Score 73; DB 2; Length 111;
Best Local Similarity 82.4%; Pred. No. 0.00011;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOYLVKIPNASLIE 17
DB 7 SHOYLVKIPNVSLID 23

RESULT 14
O71063
ID O71063 PRELIMINARY; PRT; 111 AA.
AC O71063;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026231; AAC08753.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; I.
DR Envelope protein; Fusion protein.
FT NON TER 1 111
FT NON TER 1 111
SQ SEQUENCE 111 AA; 12077 MW; 7150F267B1065AA8 CRC64;

Query Match 88.0%; Score 73; DB 2; Length 111;
Best Local Similarity 82.4%; Pred. No. 0.00011;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHOYLVKIPNASLIE 17
DB 7 SHOYLVKIPNVSLID 23

```

```
RESULT 15
071065 PRELIMINARY; PRT; 111 AA.
ID 071065
AC 071065;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liermann H., Harder T., Haas L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
    similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
    family.
DR EMBL; AF026233; AAC08755.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
FT NON_TER 1
FT NON_TER 111
SQ SEQUENCE 111 AA; 11966 MW; F26EA44A128A14B7 CRC64;

Query Match      88.0%; Score 73; DB 2; Length 111;
Best Local Similarity 82.4%; Pred. No. 0.00011;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SHQYLVKILIPNASLIE 17
   |||||:||||:
Db 7 SHQYLVKILMPNVSLID 23

Search completed: September 28, 2005, 17:55:52
Job time : 81.5 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:56:03 ; Search time 87 Seconds  
(without alignments)  
79.534 Million cell updates/sec

Title: US-10-705-819B-5

Perfect score: 89

Sequence: 1 TRSRKQTSRLKNIPVH 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1826554 seqs, 407025358 residues

Total number of hits satisfying chosen parameters: 1826554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	89	100.0	17	17 US-10-935-155A-5	Sequence 5, Appli
2	89	100.0	17	18 US-10-705-819B-5	Sequence 5, Appli
3	89	100.0	662	10 US-09-951-061A-141	Sequence 141, Appl
4	89	100.0	662	15 US-10-670-695-36	Sequence 36, Appl
5	89	100.0	662	17 US-10-935-155A-5	Sequence 27, Appl
6	89	100.0	662	18 US-10-705-819B-27	Sequence 138855,
7	45.5	51.1	226	16 US-10-437-963-138855	Sequence 349914,
8	45	50.6	154	16 US-10-425-115-349914	Sequence 30460, A
9	44.5	50.0	280	18 US-10-450-763-30460	Sequence 30644, A
10	44.5	50.0	280	18 US-10-450-763-30644	Sequence 7333, Ap
11	44	49.4	103	18 US-10-724-972A-7333	

Sequence 246424,  
Sequence 171167,  
Sequence 11286, A  
Sequence 16556, A  
Sequence 59887, A  
Sequence 14221, A  
Sequence 149975,  
Sequence 213902,  
Sequence 196724,  
Sequence 53, Appl  
Sequence 10266, A  
Sequence 56508, A  
Sequence 136326,  
Sequence 225155,  
Sequence 210104,  
Sequence 192193,  
Sequence 42, Appl  
Sequence 7, Appl  
Sequence 44, Appl  
Sequence 629, App  
Sequence 280472,  
Sequence 44240, A  
Sequence 116, App  
Sequence 2056, Ap  
Sequence 482, App  
Sequence 294295,  
Sequence 62454, A  
Sequence 66826, A  
Sequence 309154,  
Sequence 1562, Ap  
Sequence 6, Appl  
Sequence 8, Appl  
Sequence 1099, Ap  
Sequence 276230,

US-10-425-115-246424  
US-10-424-599-171167  
US-10-425-114-11286  
US-10-369-493-16556  
US-10-425-114-59887  
US-10-732-923-14221  
US-10-424-599-149975  
US-10-425-115-213902  
US-10-424-599-196724  
US-10-722-045-53  
US-10-369-493-10266  
US-10-282-122A-56508  
US-10-437-963-136326  
US-10-424-599-225155  
US-10-425-115-210104  
US-10-437-963-192193  
US-10-094-466-42  
US-10-416-314-7  
US-10-094-466-44  
US-10-719-993-629  
US-10-425-115-280472  
US-10-767-701-44240  
US-10-286-264-116  
US-10-374-780A-2056  
US-10-412-699B-482  
US-10-425-115-294295  
US-10-425-114-62454  
US-10-425-114-66826  
US-10-425-115-309154  
US-10-012-697-1562  
US-10-783-989-6  
US-10-783-989-8  
US-10-408-765A-1099  
US-10-425-115-276230

#### ALIGNMENTS

#### RESULT 1

US-10-935-155A-5  
; Sequence 5, Application US/10935155A  
; Publication No. US20050037014A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-212  
; CURRENT FILING DATE: 2004-09-08  
; PRIOR FILING DATE: 2004-09-08  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 10/705,819  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P10  
US-10-935-155A-5

```
Query Match      100.0%; Score 89; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||
Db 1 TRSRKQTSRLKNIPVH 17

RESULT 2
US-10-705-819b-5
; Sequence 5, Application US/10705819B
; Publication No. US20050186220A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-201
; CURRENT APPLICATION NUMBER: US/10/705,819B
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 09/890650
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: PCT/AU00/00070
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: AU PP8533
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v
US-10-705-819b-5

Query Match      100.0%; Score 89; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||
Db 1 TRSRKQTSRLKNIPVH 17

RESULT 3
US-09-951-061A-141
; Sequence 141, Application US/09951061A
; Publication No. US20030082204A1
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: Tartaglia, James
; APPLICANT: Taylor, Jill
; APPLICANT: Gettig, Russell
; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE
; TITLE OF INVENTION: RECOMBINANTS
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell, Boehnen, Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/951,061A
; FILING DATE: 13-SEP-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/354,138
; FILING DATE: 15-JUL-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/224,657
; FILING DATE: 16-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,962
; FILING DATE: 08-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,867
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,283
; FILING DATE: 31-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/713,967
; FILING DATE: 11-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07,666,056
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-951-061A-141

Query Match      100.0%; Score 89; DB 10; Length 662;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||
Db 64 TRSRKQTSRLKNIPVH 80

RESULT 4
US-10-670-695-36
; Sequence 36, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
```



Qy 2 RSR-KQTSRLKNIPV 16  
||: ||: ||: ||: ||: ||  
Db 110 RSKEKETNHRPRNPV 125

RESULT 8  
US-10-425-115-349914  
; Sequence 349914, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 349914  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_82296C.1.pep  
US-10-425-115-349914  
  
Query Match 50.6%; Score 45; DB 16; Length 154;  
Best Local Similarity 50.8%; Pred. No. 29;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
QY 4 RKQTSRLKKNIPVH 17  
|||:|:|:|:|  
DB 43 RKNSRHMKVSVH 56  
  
RESULT 9  
US-10-450-763-30460  
; Sequence 30460, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 30460  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (188)..(198)  
; OTHER INFORMATION: Sushii domain proteins (SCR repeat) domain identified  
; OTHER INFORMATION: by eMATRIX, accession number PF00084C, p-value=1.000e-08, raw sco  
US-10-450-763-30460  
  
Query Match 50.0%; Score 44.5; DB 18; Length 280;  
Best Local Similarity 62.5%; Pred. No. 65;  
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
  
QY 2 RSRKQTSRHR-LKNIPV 16  
|||:|:|:|:|  
DB 230 RSRRTNHRQMGRIPV 245  
  
RESULT 11  
US-10-724-972A-7333  
; Sequence 7333, Application US/10724972A  
; Publication No. US20040147734A1  
; GENERAL INFORMATION:  
; APPLICANT: Bush, David  
; APPLICANT: Doucette-Stamm, Lynn  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: PATH03-16  
; CURRENT APPLICATION NUMBER: US/10/724,972A  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: 09/450,969  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: 09/134,001  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 7544  
; SEQ ID NO 7333  
; LENGTH: 103  
; TYPE: PRT

QY 2 RSRKQTSRHR-LKNIPV 16  
|||:|:|:|:|  
DB 230 RSRRTNHRQMGRIPV 245  
  
RESULT 10  
US-10-450-763-30644  
; Sequence 30644, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 30644  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (188)..(198)  
; OTHER INFORMATION: Sushii domain proteins (SCR repeat) domain identified  
; OTHER INFORMATION: by eMATRIX, accession number PF00084C, p-value=1.000e-08, raw sco  
US-10-450-763-30644  
  
Query Match 50.0%; Score 44.5; DB 18; Length 280;  
Best Local Similarity 62.5%; Pred. No. 65;  
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
  
QY 2 RSRKQTSRHR-LKNIPV 16  
|||:|:|:|:|  
DB 230 RSRRTNHRQMGRIPV 245  
  
RESULT 11  
US-10-724-972A-7333  
; Sequence 7333, Application US/10724972A  
; Publication No. US20040147734A1  
; GENERAL INFORMATION:  
; APPLICANT: Bush, David  
; APPLICANT: Doucette-Stamm, Lynn  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: PATH03-16  
; CURRENT APPLICATION NUMBER: US/10/724,972A  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: 09/450,969  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: 09/134,001  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 7544  
; SEQ ID NO 7333  
; LENGTH: 103  
; TYPE: PRT

; ORGANISM: S.epidermidis  
 US-10-724-972A-7333

Query Match 49.4%; Score 44; DB 18; Length 103;  
 Best Local Similarity 69.2%; Pred. No. 28;  
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 KQTSRLKKNIPVH 17  
 Db 81 KQSDRLKEIPVH 93

RESULT 12  
 US-10-425-115-246424  
 ; Sequence 246424, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 246424  
 ; LENGTH: 143  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_15631C.1.pep  
 US-10-425-115-246424

Query Match 49.4%; Score 44; DB 16; Length 143;  
 Best Local Similarity 35.3%; Pred. No. 39;  
 Matches 6; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKKNIPVH 17  
 Db 84 TKQRTYRLAIVPH 100

RESULT 13  
 US-10-424-599-171167  
 ; Sequence 171167, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 171167  
 ; LENGTH: 197  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_125579C.1.pep  
 US-10-424-599-171167

Query Match 48.9%; Score 43.5; DB 15; Length 197;  
 Best Local Similarity 41.7%; Pred. No. 65;  
 Matches 10; Conservative 4; Mismatches 3; Indels 7; Gaps 1;

Qy 1 TRSRKQTSRLKKN-----IPVH 17  
 Db 84 KQSDRLKEIPVH 93

; ORGANISM: S.epidermidis  
 US-10-724-972A-7333

Query Match 48.3%; Score 43; DB 15; Length 209;  
 Best Local Similarity 46.2%; Pred. No. 83;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KQTSRLKKNIPVH 17  
 Db 34 KQLGNRMKDLPIH 46

RESULT 14  
 US-10-425-114-41286  
 ; Sequence 41286, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 41286  
 ; LENGTH: 200  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700572887\_FLI.pep  
 US-10-425-114-41286

Query Match 48.3%; Score 43; DB 15; Length 200;  
 Best Local Similarity 50.0%; Pred. No. 80;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RSRKQTSRLKKNIP 15  
 Db 51 RRRQRRLRRVP 64

RESULT 15  
 US-10-369-493-16556  
 ; Sequence 16556, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 16556  
 ; LENGTH: 209  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus thuringiensis  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(209)  
 ; OTHER INFORMATION: unsure at all xaa locations  
 US-10-369-493-16556

Query Match 48.3%; Score 43; DB 15; Length 209;  
 Best Local Similarity 46.2%; Pred. No. 83;  
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 5 KQTSRLKKNIPVH 17  
 Db 34 KQLGNRMKDLPIH 46

Search completed: September 28, 2005, 18:26:02  
Job time : 88 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:41:21 ; Search time 23.5 Seconds  
(without alignments)  
54.001 Million cell updates/sec

Title: US-10-705-819B-5  
Perfect score: 89  
Sequence: 1 TRSRKQTSRLKNIPVH 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pap.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	4	US-09-890-650-5
2	89	100.0	662	1	Sequence 5, Appli
3	89	100.0	662	3	Sequence 88, Appl
4	89	100.0	662	4	Sequence 27, Appl
5	44	49.4	103	3	US-09-890-650-27
6	44	49.4	113	4	US-09-134-001C-5577
7	44	49.4	245	4	US-09-270-767-41615
8	42	47.2	186	5	Sequence 58061, A
9	42	47.2	937	3	PCT-US91-08177-15
10	42	47.2	944	4	US-09-005-180A-4
11	42	47.2	1720	2	US-09-538-092-452
12	40	44.9	238	4	Sequence 452, App
13	40	44.9	238	4	Sequence 12, Appl
14	40	44.9	758	4	US-09-830-433A-16
15	39	43.8	63	4	US-09-830-433A-63
16	39	43.8	293	4	US-09-252-991A-23363
17	39	43.8	488	4	US-09-634-238-272
18	39	43.8	613	4	Sequence 82, Appl
19	39	43.8	613	4	US-09-800-729-98
20	39	43.8	627	4	Sequence 98, Appl
21	38	42.7	60	4	Sequence 88, Appl
22	38	42.7	107	4	Sequence 6304, Ap
23	38	42.7	122	4	Sequence 10083, A
24	38	42.7	125	4	Sequence 23644, A
25	38	42.7	191	4	US-09-270-767-60860
26	38	42.7	341	4	Sequence 60860, A
27	38	42.7	469	4	Sequence 16895, A
					Sequence 26021, A
					Sequence 45360, A

28 38 42.7 471 4 US-09-252-991A-30899 Sequence 30899, A  
29 38 42.7 489 4 US-03-583-110-4914 Sequence 4914, Ap  
30 38 42.7 497 4 US-09-107-433-4571 Sequence 4571, Ap  
31 38 42.7 572 4 US-09-248-796A-19066 Sequence 19066, A  
32 38 42.7 618 4 US-09-252-991A-22418 Sequence 22418, A  
33 38 42.7 769 4 US-09-949-016-8149 Sequence 8149, Ap  
34 38 42.7 784 4 US-09-252-991A-30377 Sequence 30377, A  
35 38 42.7 806 4 US-09-949-016-11459 Sequence 11459, A  
36 38 42.7 953 4 US-09-252-991A-27230 Sequence 27230, A  
37 38 42.7 1073 4 US-09-949-016-9771 Sequence 9771, A  
38 38 42.7 1083 4 US-03-394-272-11 Sequence 11, Appli  
39 38 42.7 1104 2 US-08-327-832-5 Sequence 5, Appli  
40 38 42.7 1104 2 US-08-828-584-5 Sequence 5, Appli  
41 37 41.6 51 4 US-09-205-258-430 Sequence 430, App  
42 37 41.6 102 4 US-09-583-110-3604 Sequence 3604, App  
43 37 41.6 104 4 US-09-270-767-57705 Sequence 57705, A  
44 37 41.6 106 4 US-09-107-433-4126 Sequence 4126, Ap  
45 37 41.6 272 4 US-09-710-279-2290 Sequence 2290, Ap

## ALIGNMENTS

## RESULT 1

US-09-890-650-5  
; Sequence 5, Application US/09890650  
; Patent No. 6685947  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, DAVID CHARLES  
; APPLICANT: SOURAVI, GHOSH  
; APPLICANT: WALKER, JOHN  
; TITLE OF INVENTION: T HELPER CELL EPITOPES  
; FILE REFERENCE: 47-152  
; CURRENT APPLICATION NUMBER: US/09/890,650  
; CURRENT FILING DATE: 2001-11-05  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: canine distemper virus  
US-09-890-650-5

Query Match 100.0%; Score 89; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.3e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17  
Db 1 TRSRKQTSRLKNIPVH 17

## RESULT 2

US-08-224-657-88  
; Sequence 88, Application US/08224657  
; Patent No. 5756102  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: Tartaglia, James  
; APPLICANT: Taylor, Jill  
; TITLE OF INVENTION: FOXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
; TITLE OF INVENTION: RECOMBINANTS  
; NUMBER OF SEQUENCES: 122  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/224,657  
FILING DATE: 06-APR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2550  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 662 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-224-657-88  
Query Match 100.0%; Score 89; DB 1; Length 662;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1 TRSRKQTSRLKNIPVH 17  
Db 64 TRSRKQTSRLKNIPVH 80  
RESULT 3  
US-09-354-138-88  
Sequence 88, Application US/09354138  
Patent No. 6309647  
GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
APPLICANT: Tartaglia, James  
APPLICANT: Taylor, Jill  
APPLICANT: Gettig, Russell  
TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)  
TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE  
TITLE OF INVENTION: RECOMBINANTS  
NUMBER OF SEQUENCES: 139  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford, P.C.  
STREET: 530 Fifth Avenue, 25th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/354,138  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/472,379  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: US 08/416,646  
FILING DATE: 05-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/224,657  
FILING DATE: 16-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/073,962

FILING DATE: 08-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/776,867  
FILING DATE: 23-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/621,614  
FILING DATE: 30-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/938,283  
FILING DATE: 31-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/105,483  
FILING DATE: 12-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/713,967  
FILING DATE: 11-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07,666,056  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2860  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 88:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 662 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-09-354-138-88  
Query Match 100.0%; Score 89; DB 3; Length 662;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1 TRSRKQTSRLKNIPVH 17  
Db 64 TRSRKQTSRLKNIPVH 80  
RESULT 4  
US-09-890-650-27  
Sequence 27, Application US/09890650  
Patent No. 6685947  
GENERAL INFORMATION:  
APPLICANT: JACKSON, DAVID CHARLES  
APPLICANT: SOURAVI, GHOSH  
APPLICANT: WALKER, JOHN  
TITLE OF INVENTION: T HELPER CELL EPITOPES  
FILE REFERENCE: 47-152  
CURRENT APPLICATION NUMBER: US/09/890,650  
CURRENT FILING DATE: 2001-11-05  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 27  
LENGTH: 662  
TYPE: PRT  
ORGANISM: canine distemper virus  
US-09-890-650-27  
Query Match 100.0%; Score 89; DB 4; Length 662;  
Best Local Similarity 100.0%; Pred. No. 1.7e-06; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;  
QY 1 TRSRKQTSRLKNIPVH 17

```

Db      64 TRSRKQTSRLKNIPVH 80
|||||
RESULT 5
US-09-134-001C-5577
; Sequence 5577, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5577
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5577
Query Match          49.4%; Score 44; DB 3; Length 103;
Best Local Similarity 69.2%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY    5 KQTSRLKNIPVH 17
      | |||||
DB    81 KNQSDRLKEIPVH 93

RESULT 6
US-09-270-767-41615
; Sequence 41615, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41615
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41615
Query Match          49.4%; Score 44; DB 4; Length 113;
Best Local Similarity 56.2%; Pred. No. 6.4;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY    2 RSRKQTSRLKNIPVH 17
      | ||| :|||
DB   23 QRNRATGNRLKCAPVH 38

RESULT 7
US-09-270-767-58061
; Sequence 58061, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

```

```
; Patent No. 612446
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN VPS35/MEM3-RELATED PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,180A
; FILING DATE: Filed January 8, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0457 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 937 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 854543
; US-09-005-180A-4

Query Match 47.2%; Score 42; DB 3; Length 937;
Best Local Similarity 47.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
Db 904 TASGSKLLHQLKIPIH 920

RESULT 10
US-09-538-092-452
; Sequence 452, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormat Version 0.9
; SEQ ID NO 452
; LENGTH: 944

; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YJL154C
US-09-538-092-452

Query Match 47.2%; Score 42; DB 4; Length 944;
Best Local Similarity 47.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
Db 904 TASGSKLLHQLKIPIH 920

RESULT 11
US-08-477-451-12
; Sequence 12, Application US/08477451
; Patent No. 5928865
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; TITLE OF INVENTION: Helicobacter Pylori Cagi Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1720 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-451-12

Query Match 47.2%; Score 42; DB 2; Length 1720;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 TSHRLKNIPV 16
Db 122 TSKRLKNIFI 131

RESULT 12
US-09-830-433A-16
; Sequence 16, Application US/09830433A
; Patent No. 6835384
; GENERAL INFORMATION:
; APPLICANT: AUJAME et al.
; TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic
```



; TITLE OF INVENTION: strains of the Neisseria genus  
; FILE REFERENCE: P07180US00/BAS  
; CURRENT APPLICATION NUMBER: US/09/830,433A  
; CURRENT FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: FR 98 13 693  
; PRIOR FILING DATE: 1998-10-30  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 238  
; TYPE: PRT  
; ORGANISM: Neisseria meningitidis  
US-09-830-433A-16

Query Match 44.9%; Score 40; DB 4; Length 238;  
Best Local Similarity 52.9%; Pred. No. 65;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKKNIPVH 17  
Db 144 TRSLKATVERLENQKLH 160

RESULT 13  
US-09-830-433A-63  
; Sequence 63, Application US/09830433A  
; Patent No. 6835384  
; GENERAL INFORMATION:

; APPLICANT: AUJAME et al.  
; TITLE OF INVENTION: Nucleic acids and polypeptides specific for pathogenic  
; FILE REFERENCE: P07180US00/BAS  
; CURRENT APPLICATION NUMBER: US/09/830,433A

; CURRENT FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: FR 98 13 693  
; PRIOR FILING DATE: 1998-10-30  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 63  
; LENGTH: 238  
; TYPE: PRT

; ORGANISM: Neisseria meningitidis  
US-09-830-433A-63

Query Match 44.9%; Score 40; DB 4; Length 238;  
Best Local Similarity 52.9%; Pred. No. 65;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKKNIPVH 17  
Db 144 TRSLKATVERLENQKLH 160

RESULT 14  
US-09-252-991A-29974  
; Sequence 29974, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29974  
; LENGTH: 758  
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-29974

Query Match 44.9%; Score 40; DB 4; Length 758;  
Best Local Similarity 53.3%; Pred. No. 2.3e+02;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 RSRKQTSRLKKNIPV 16  
Db 41 RGRKSTRRRWANLPV 55

RESULT 15

US-09-248-796A-21747  
; Sequence 21747, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 21747  
; LENGTH: 63  
; TYPE: PRT

; ORGANISM: Candida albicans  
US-09-248-796A-21747

Query Match 43.8%; Score 39; DB 4; Length 63;  
Best Local Similarity 58.3%; Pred. No. 23;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RKQTSRLKKNIP 15  
Db 29 KKQTNNEKNIP 40

Search completed: September 28, 2005, 17:58:44  
Job time : 23.5 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:39:51 ; Search time 80.5 Seconds  
(without alignments)  
108.141 Million cell updates/sec

Title: US-10-705-819B-5

Perfect score: 89

Sequence: 1 TRSRKQTSRLKNIPVH 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	89	100.0	662	1	VGLF_CDVO
2	89	100.0	662	2	O89327
3	89	100.0	662	2	O9DX22
4	84	94.4	662	2	O91KN3
5	69	77.5	662	2	O6SVY0
6	58	65.2	266	2	O6SVY0
7	55	61.8	266	2	O55336
8	55	61.8	662	2	O68PH9
9	55	61.8	662	2	O9YKL7
10	54	60.7	266	2	O55334
11	54	60.7	266	2	O55335
12	54	60.7	662	2	O6T1B5
13	54	60.7	662	2	O6TV24
14	54	60.7	662	2	O6TV32
15	46	51.7	266	2	O55333
16	46	51.7	657	2	O6MMP1
17	45.5	51.1	226	2	O7X059
18	44	49.4	98	2	O8CTK2
19	44	49.4	416	2	O8SVN1
20	43	48.3	172	2	O43025
21	43	48.3	203	2	O81EF1
22	43	48.3	701	2	O7S6B7
23	42.5	47.8	105	2	O8AWZ7
24	42	47.2	31	2	O8KYI9
25	42	47.2	31	2	O6EZK4
26	42	47.2	186	1	VNA2_BRSVA
27	42	47.2	186	2	O77KZ7
28	42	47.2	186	2	O9YS23
29	42	47.2	334	2	O86LV3
30	42	47.2	648	2	O83LL7
31	42	47.2	726	1	ETK_ECO27

32 42 47.2 726 1 ETK\_ECO57  
33 42 47.2 726 1 ETK\_ECOL1  
34 42 47.2 934 2 Q924X9  
35 42 47.2 944 1 VP35\_YEAST  
36 42 47.2 1201 2 Q7R4Q5  
37 42 47.2 2356 2 Q7R1F7  
38 41 46.1 66 2 Q74VX3  
39 41 46.1 160 2 Q7X6Z6  
40 41 46.1 186 1 VNA2\_ORSW  
41 41 46.1 186 2 Q8WVP5  
42 41 46.1 257 2 Q7Q8P2  
43 41 46.1 270 2 Q9FFC8  
44 41 46.1 298 2 Q9XXC3  
45 41 46.1 403 2 Q93WJ9

Q8xc28 escherichia  
P38l34 escherichia  
Q924x9 mus musculus  
P34l10 saccharomyc  
O7r4q5 giardia lam  
O7rif7 plasmodium  
Q74vx3 yersinia pe  
Q7x6z6 oryza sativ  
Q84l32 ovine respi  
Q8wvp5 homo sapien  
Q7q8p2 anopheles g  
O9ffc8 arabidopsis  
Q8xxc3 caenorhabdi  
Q93wj9 arabidopsis

## ALIGNMENTS

## RESULT 1

VGLF\_CDVO  
ID VGLF\_CDVO STANDARD; PRT; 662 AA.  
AC P12569; Q65991;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DB Fusion glycoprotein precursor [Contains: Fusion glycoprotein F2;  
DE Fusion glycoprotein F1].  
GN Name=F;  
OS Canine distemper virus (strain Onderstepoort) (CDV).  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses;  
OX NCBI\_TaxID=11233;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88129050; PubMed=3433924; DOI=10.1016/0168-1702(87)90009-8;  
RA Barrett T., Clarke D.K., Evans S.A., Rima B.K.;  
RT "The nucleotide sequence of the gene encoding the F protein of canine  
distemper virus: a comparison of the deduced amino acid sequence with  
other paramyxoviruses.";  
RL Virus Res. 8:373-386(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93227696; PubMed=8470428; DOI=10.1016/0264-410X(93)90285-6;  
RA Wild T.F., Bernard A., Spehner D., Villevall D., Drillean R.;  
RT "Vaccination of mice against canine distemper virus-induced  
encephalitis with vaccinia virus recombinants encoding measles or  
canine distemper virus antigens.";  
RL Vaccine 11:438-444(1993).  
CC -!- FUNCTION: This protein directs fusion of viral and cellular  
membranes.  
CC -!- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked.  
CC -!- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; M21849; AAA42878.1; -;  
DR EMBL; X65509; CAA46481.1; -;  
DR PIR; J80321; VGNZCD.  
DR PIR; S21382; S21382.  
DR HSSP; P04849; 1SVF.  
DR InterPro; IPR000776; Fusion gly.  
DR Pfam; PF00523; Fusion gly; 1.  
KW Envelope protein; Fusion protein; Glycoprotein; Signal; Transmembrane.  
FT SIGNAL 1 ? 662 Fusion glycoprotein F0.  
FT CHAIN ? 662

```

FT CHAIN ? 224 Fusion glycoprotein F2.
FT CHAIN 225 662 Fusion glycoprotein F1.
FT TRANSMEM 606 629 Potential.
FT DISULFID 180 307 Linkage between F2 and F1 (Potential).
FT CARBOHYD 62 62 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 141 141 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 173 173 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 179 179 N-linked (GlcNAc...) (Potential).
FT CONFLICT 3 3 R -> K (in Ref. 2).
FT CONFLICT 140 140 D -> N (in Ref. 2).
FT CONFLICT 152 152 N -> S (in Ref. 2).
FT CONFLICT 171 171 I -> M (in Ref. 2).
FT CONFLICT 174 174 A -> V (in Ref. 2).
FT CONFLICT 662 662 L -> H (in Ref. 2).
SQ SEQUENCE 662 AA; 72970 MW; FB2C81C9797805F0 CRC64;

Query Match 100.0%; Score 89; DB 1; Length 662;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
DB 64 TRSRKQTSRLKNIPVH 80

RESULT 2
O89327 PRELIMINARY; PRT; 662 AA.
AC O89327;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93174978; PubMed=8438593;
RA Sidhu M.S., Hsuar W., Cook S.D., Dowling P.C., Udem S.A.;
RT "Canine distemper terminal and intergenic non-protein coding
nucleotide sequences: completion of the entire CDV genome sequence.";
RL Virology 193:66-72(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Sidhu M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF014953; AAC26994.1; -.
DR HSSP; P04849; ISVF.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72951 MW; 80B144C6B9801898 CRC64;

Query Match 100.0%; Score 89; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
DB 64 TRSRKQTSRLKNIPVH 80

RESULT 3
O9DXZ2 PRELIMINARY; PRT; 662 AA.
AC O9DXZ2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20499096; PubMed=11044118;
RX DOI=10.1128/JVI.74.22.10737-10744.2000;
RA Gassen U., Collins F.M., Duprex W.P., Rima B.K.;
RT "Establishment of a rescue system for canine distemper virus.";
RL J. Virol. 74:10737-10744 (2000).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF305419; AAG30919.1; -.
DR HSSP; P04849; ISVF.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72988 MW; 9C5C1398C9AE7B4C CRC64;

Query Match 100.0%; Score 89; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
DB 64 TRSRKQTSRLKNIPVH 80

```

```

ID O9DXZ2 PRELIMINARY; PRT; 662 AA.
AC O9DXZ2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20499096; PubMed=11044118;
RX DOI=10.1128/JVI.74.22.10737-10744.2000;
RA Gassen U., Collins F.M., Duprex W.P., Rima B.K.;
RT "Establishment of a rescue system for canine distemper virus.";
RL J. Virol. 74:10737-10744 (2000).
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF305419; AAG30919.1; -.
DR HSSP; P04849; ISVF.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0006948; P:Viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72988 MW; 9C5C1398C9AE7B4C CRC64;

Query Match 100.0%; Score 89; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
DB 64 TRSRKQTSRLKNIPVH 80

RESULT 4
Q91KN3 PRELIMINARY; PRT; 662 AA.
ID Q91KN3;
AC Q91KN3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus (strain Onderstepoort) (CDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11233;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Onderstepoort;
RX MEDLINE=21306344; PubMed=11413309;
RX DOI=10.1128/JVI.75.14.6418-6427.2001;
RA von Messling V., Zimmer G., Herrler G., Haas L., Cattaneo R.;
RT "The hemagglutinin of canine distemper virus determines tropism and
cytopathogenicity.";
RL J. Virol. 75:6418-6427 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Onderstepoort;
RA von Messling V.A., Zimmer G., Herrler G., Haas L., Cattaneo R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
family.
DR EMBL; AF378705; AAK54668.1; -.
DR HSSP; P04849; ISVF.
DR GO; GO:0019031; C:Viral envelope; IEA.

```

DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR Pfam; PF00523; Fusion gly; 1.  
 KW Envelope protein; Fusion protein.  
 SQ SEQUENCE 662 AA; 72898 MW; C6A104A96BBF8A0 CRC64;

Query Match 94.4%; Score 84; DB 2; Length 662;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-06;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RSRKQTSRLKNIPVH 17  
 ||||| ||||| |||||  
 Db 65 RSRKQTSRLKNIPVH 80

RESULT 5  
 ID Q6SY0 PRELIMINARY; PRT; 662 AA.  
 AC Q6SY0;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Fusion protein F.  
 OS Canine distemper virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lednický J.A., Meehan T.P., Sarich N.A.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;  
 RA Lednický J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,  
 RA Sarich N.A., Witteky K.E., Braid M.D., Pedrak C., Houde C.M.;  
 RT "Effective primary isolation of wild-type Canine distemper virus in  
 MDCK, MV1 Lu and Vero cells without nucleotide sequence changes within  
 RT the entire haemagglutinin protein gene and in subgenomic sections of  
 RT the fusion and phosphoprotein genes";  
 RL J. Virol. Methods 118:147-157(2004).  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 family.  
 DR EMBL; AY445077; AARI6539.1; -  
 DR EMBL; AY542312; AAS48411.1; -  
 DR EMBL; AY466011; AAR30102.1; -  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro; IPR000776; Fusion gly.  
 DR InterPro; IPR009050; Globin\_like.  
 DR Pfam; PF00523; Fusion gly; 1.  
 KW Envelope protein; Fusion protein.  
 SQ SEQUENCE 662 AA; 72736 MW; 2DD1B45291D3D315 CRC64;

Query Match 77.5%; Score 69; DB 2; Length 662;  
 Best Local Similarity 81.2%; Pred. No. 0.0028;  
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RSRKQTSRLKNIPVH 17  
 ||||| ||||| |||||  
 Db 65 RSRKQTSRLKNIPVH 80

RESULT 6  
 ID O55338 PRELIMINARY; PRT; 266 AA.  
 AC O55338;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Fusion protein (fragment).  
 OS Canine distemper virus.

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liemann H., Harder T., Haas L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 family.  
 DR EMBL; AF026244; AAB88269.1; -  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro; IPR000776; Fusion gly.  
 DR InterPro; IPR009050; Globin\_like.  
 DR Pfam; PF00523; Fusion gly; 1.  
 KW Envelope protein; Fusion protein.  
 FT NON TER 266  
 SQ SEQUENCE 266 AA; 29389 MW; BA762A02F1F39DEC CRC64;

Query Match 65.2%; Score 58; DB 2; Length 266;  
 Best Local Similarity 70.6%; Pred. No. 0.085;  
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17  
 ||||| ||||| |||||  
 Db 64 TRSRKQTSRLKNIPVH 80

RESULT 7  
 ID O55336 PRELIMINARY; PRT; 266 AA.  
 AC O55336;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Fusion protein (fragment).  
 OS Canine distemper virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liemann H., Harder T., Haas L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 family.  
 DR EMBL; AF026242; AAB88267.1; -  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro; IPR000776; Fusion gly.  
 DR InterPro; IPR009050; Globin\_like.  
 DR Pfam; PF00523; Fusion gly; 1.  
 KW Envelope protein; Fusion protein.  
 FT NON TER 266  
 SQ SEQUENCE 266 AA; 29135 MW; 16A6FEA46AEC1DAC CRC64;

Query Match 61.8%; Score 55; DB 2; Length 266;  
 Best Local Similarity 64.7%; Pred. No. 0.29;  
 Matches 11; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17  
 ||||| ||||| |||||  
 Db 64 TRSRKQTSRLKNIPVH 80

RESULT 8  
 Q68PH9 PRELIMINARY; PRT; 662 AA.  
 ID Q68PH9  
 AC Q68PH9;

```

DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fusion protein.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Lednický J.A., Meehan T.P., Sarich N.A., Witecki K.E.;
RT "Full genomic sequence of Canine distemper virus 01-2689."
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AV649446; AAT94552.1; -.
DR InterPro; IPR000776; Fusion_gly.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72565 MW; 656A8757BFD763AA CRC64;

Query Match 61.8%; Score 55; DB 2; Length 662;
Best Local Similarity 68.8%; Pred. No. 0.84;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RSRKQTSRLKNIPVH 17
DB 65 RSCQASHRSNDNIPAH 80

RESULT 9
QYQYKL7 PRELIMINARY; PRT; 662 AA.
ID QYQYKL7
AC QYQYKL7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 03-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fusion protein.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Cherpillod P., Beck K., Zurbriggen A., Wittek R.;
RT "Sequence analysis and expression of the attachment and fusion
RT proteins of canine distemper virus wild-type strain A75/17."
RL J. Virol. 73:2263-2269(1999).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=A75/17;
CC MEDLINE=99139009; PubMed=9971809;
RA Cherpillod P., Beck K., Zurbriggen A., Wittek R.;
RT "Sequence analysis and expression of the attachment and fusion
RT proteins of canine distemper virus wild-type strain A75/17."
RL J. Virol. 73:2263-2269(1999).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=A75/17;
CC MEDLINE=99139009; PubMed=9971809;
RA Wiedeker C., Howley P., Zurbriggen A., Wittek R.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF112188; AAD18007.1; -.
DR EMBL; AF164967; AAD49702.1; -.
DR HSP; P04849; 1SVF.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72537 MW; 68F992DCBA51F0BA CRC64;

Query Match 61.8%; Score 55; DB 2; Length 662;

Best Local Similarity 68.8%; Pred. No. 0.84;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 RSRKQTSRLKNIPVH 17
DB 65 RSCQASHRSNDNIPAH 80

RESULT 10
QYQYKL7 PRELIMINARY; PRT; 266 AA.
ID QYQYKL7
AC QYQYKL7;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026240; AAB88265.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.
KW Envelope protein; Fusion protein.
FT NON TER 266
SQ SEQUENCE 266 AA; 29268 MW; 68AD69FB27604FA5 CRC64;

Query Match 60.7%; Score 54; DB 2; Length 266;
Best Local Similarity 64.7%; Pred. No. 0.43;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17
DB 64 TRSCKQASHRSNDNIPAH 80

RESULT 11
QYQYKL7 PRELIMINARY; PRT; 266 AA.
ID QYQYKL7
AC QYQYKL7;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Fusion protein (Fragment).
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RA Liemann H., Harder T., Haas L.;
RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AF026241; AAB88266.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion_gly.
DR InterPro; IPR009050; Globin_like.
DR Pfam; PF00523; Fusion_gly; 1.

```

```

KW Envelope protein; Fusion protein.
FT NON TER 266
SQ SEQUENCE 266 AA; 29191 MW; 7A80B29F09E872EE CRC64;

Query Match 60.7%; Score 54; DB 2; Length 266;
Best Local Similarity 64.7%; Pred. No. 0.43;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||:|:|:|
Db 64 TRSCQASYSRSDNIPAH 80

RESULT 12
Q6TVB5 PRELIMINARY; PRT; 662 AA.
ID Q6TVB5
AC Q6TVB5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
SQ SEQUENCE FROM N.A.
RX PubMed=15081610; DOI=10.1016/j.jviromet.2004.02.004;
RA Lednicki J.A., Meehan T.P., Kinsel M.J., Dubach J., Hungerford L.L.,
RA Sarich N.A., Witteck K.E., Braid M.D., Pedrak C., Houde C.M.;
RT "Effective primary isolation of wild-type Canine distemper virus in
RT MDCK, Mv1 Lu and Vero cells without nucleotide sequence changes within
RT the entire haemagglutinin protein gene and in subgenomic sections of
RT the fusion and phosphoprotein genes.";
RL J. Virol. Methods 118:147-157(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Lednicki J.A., Meehan T.P., Sarich N.A.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY443350; AAR32272.1; -.
DR EMBL; AY395984; AAQ96334.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR Pfam; PF00523; Fusion_gly; I.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72544 MW; E7B4A3BD7D734B77 CRC64;

Query Match 60.7%; Score 54; DB 2; Length 662;
Best Local Similarity 68.8%; Pred. No. 1.3;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RSRKQTSRLKNIPVH 17
   |||||:|:|:|
Db 65 RSCQTSYSRSDNIPAH 80

RESULT 13
Q6TV24 PRELIMINARY; PRT; 662 AA.
ID Q6TV24
AC Q6TV24;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;

```

```

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=5804P;
RX MEDLINE=22972557; PubMed=14610181;
RX DOI=10.1128/JVI.77.23.12579-12591.2003;
RA Von Messling V., Springfield C., Devaux P., Cattaneo R.;
RT "A ferret model of canine distemper virus virulence and
RT immunosuppression.";
RL J. Virol. 77:12579-12591(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5804P;
RA von Messling V., Springfield C., Devaux P., Cattaneo R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY386316; AAQ96307.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR InterPro; IPR009050; Globin-like.
DR Pfam; PF00523; Fusion_gly; I.
KW Envelope protein; Fusion protein.
SQ SEQUENCE 662 AA; 72733 MW; 43352909C3199DCB CRC64;

Query Match 60.7%; Score 54; DB 2; Length 662;
Best Local Similarity 64.7%; Pred. No. 1.3;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||:|:|:|
Db 64 TRSCQASYSRSDNIPAH 80

RESULT 14
Q6TV32 PRELIMINARY; PRT; 662 AA.
ID Q6TV32
AC Q6TV32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fusion protein F.
OS Canine distemper virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=11232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5804;
RX MEDLINE=22972557; PubMed=14610181;
RX DOI=10.1128/JVI.77.23.12579-12591.2003;
RA Von Messling V., Springfield C., Devaux P., Cattaneo R.;
RT "A ferret model of canine distemper virus virulence and
RT immunosuppression.";
RL J. Virol. 77:12579-12591(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=5804;
RA von Messling V., Springfield C., Devaux P., Cattaneo R.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By
CC similarity).
CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein
CC family.
DR EMBL; AY386315; AAQ96299.1; -.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.
DR InterPro; IPR000776; Fusion gly.
DR InterPro; IPR009050; Globin-like.
DR Pfam; PF00523; Fusion_gly; I.
KW Envelope protein; Fusion protein.

```

SQ SEQUENCE 662 AA; 72705 MW; 91F71B34F84DAC0C CRC64;  
 Query Match 60.7%; Score 54; DB 2; Length 662;  
 Best Local Similarity 64.7%; Pred. No. 1.3;  
 Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TRSRKQTSHELKNIPVH 17  
 ||||| :|||  
 Db 64 TRSCQASYRSDNIPAH 80

RESULT 15

OS5333 PRELIMINARY; PRT; 266 AA.  
 AC OS5333;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Fusion protein (Fragment).  
 OS Canine distemper virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
 OX NCBI\_TaxID=11232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liemann H., Harder T., Haas L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: Heterodimer of F1 and F2; disulfide-linked (By  
 similarity).  
 CC -1- SIMILARITY: Belongs to the paramyxoviruses fusion glycoprotein  
 family.  
 DR EMBL; AF026239; AAB88264.1; -.  
 DR GO; GO:0019031; C:viral envelope; IEA.  
 DR GO; GO:0006948; P:viral-induced cell-cell fusion; IEA.  
 DR InterPro; IPR000776; Fusion gly.  
 DR InterPro; IPR009050; Globin-like.  
 DR Pfam; PF00523; Fusion gly; I.  
 KW Envelope protein; Fusion protein.  
 FT NON\_TER 266  
 SQ SEQUENCE 266 AA; 29164 MW; CBC813FC0A366375 CRC64;

Query Match 51.7%; Score 46; DB 2; Length 266;  
 Best Local Similarity 66.7%; Pred. No. 11;  
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TRSRKQTSHELKNIP 15  
 ||||| :|||  
 Db 64 TRSCQASYRSDNIP 78

Search completed: September 28, 2005, 17:55:51  
 Job time : 82.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:40:37 ; Search time 17 Seconds  
(without alignments)  
96,217 Million cell updates/sec

Title: US-10-705-819B-5  
Perfect score: 89  
Sequence: 1 TRSRKQTSRLKNIPVH 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79:\*  
1: Pirl.\*  
2: Pirl.\*  
3: Pirl.\*  
4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	662	1 VGNZCD	cell fusion glycop
2	89	100.0	662	2 S21382	cell fusion protei
3	43	48.3	172	2 T40291	hypothetical prote
4	42	47.2	186	1 WNNZBA	matrix glycoprotei
5	42	47.2	726	2 A90771	hypothetical prote
6	42	47.2	726	2 E85633	hypothetical prote
7	42	47.2	726	2 C64839	probable ATPase yc
8	42	47.2	944	2 S56936	vacuolar protein-s
9	41	46.1	298	2 T27090	hypothetical prote
10	41	46.1	464	2 T01324	hypothetical prote
11	40	44.9	106	2 S56222	probable membrane
12	40	44.9	194	1 WNNZ22	envelope-associate
13	40	44.9	686	2 G82448	sensor histidine k
14	40	44.9	2264	1 GNVV7B	genome polypeptin
15	39	43.8	209	2 C96999	hypothetical prote
16	39	43.8	275	2 B48600	RNA-binding protei
17	39	43.8	294	2 A49688	lactose-binding le
18	39	43.8	522	2 H72268	conserved hypotet
19	39	43.8	536	2 H72267	conserved hypotet
20	39	43.8	584	2 T09113	response regulator
21	39	43.8	659	2 T11641	hypothetical prote
22	39	43.8	752	2 S64750	probable ATP-depen
23	39	43.8	1295	2 S60179	pol polyprotein ho
24	38.5	43.3	334	2 A10795	probable receptor/
25	38	42.7	112	2 T01515	hypothetical prote
26	38	42.7	127	2 A84282	hypothetical prote
27	38	42.7	170	2 S52754	ESs1 protein - yea
28	38	42.7	246	2 F91238	PTS system, fructo
29	38	42.7	246	2 B86086	PTS system, fructo

30 38 42.7 306 2 A88040 protein F47F6.1 [i  
31 38 42.7 347 2 G95402 hypothetical prote  
32 38 42.7 359 2 H65201 pts system, fructo  
33 38 42.7 365 2 T20958 hypothetical prote  
34 38 42.7 404 2 T27161 hypothetical prote  
35 38 42.7 441 2 E90051 hypothetical prote  
36 38 42.7 464 2 A88000 protein Y54E5A.3 [i  
37 38 42.7 1083 2 T04062 sucrose-phosphate  
38 38 42.7 1104 2 T38869 transcription fact  
39 38 42.7 4930 2 E69679 polyketide synthet  
40 37.5 42.1 958 2 T40466 hypothetical prote  
41 37.5 42.1 983 2 B49284 immediate-early pr  
42 37.5 42.1 1078 2 T44232 hypothetical prote  
43 37 41.6 39 1 S28546 protamine 1 - Japa  
44 37 41.6 99 2 G89838 hypothetical prote  
45 37 41.6 104 2 F98096 hypothetical prote

#### ALIGNMENTS

##### RESULT 1

VGNZCD

cell fusion glycoprotein precursor - canine distemper virus

N;Contains: fusion protein F1; fusion protein F2

C;Species: canine distemper virus

C;Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004

C;Accession: JS0321

R;Barrett, T.; Clarke, D.K.; Evans, S.A.; Rima, B.K.

Virus Res. 8, 373-386, 1987

A;Title: The nucleotide sequence of the gene encoding the F protein of canine distemper

A;Reference number: JS0321; MUID:88129050; PMID:3433924

A;Accession: JS0321

A;Molecule type: mRNA

A;Residues: 1-662 <BAR>

A;Cross-references: UNIPROT:P12569; GB:M21849; NID:G323241; PIDN:AAA42878.1; PID:G323242

C;Gene(s):

A;Gene: F

C;Superfamily: parainfluenza virus cell fusion protein

C;Keywords: glycoprotein; membrane fusion; transmembrane protein

F;1-135/Domain: signal sequence #status predicted <SIG>

F;136-224/Product: cell fusion glycoprotein F2 #status predicted <F2P>

F;225-662/Product: cell fusion glycoprotein F1 #status predicted <F1P>

F;606-629/Domain: transmembrane #status predicted <MEM>

F;62,141,173,179,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 89; DB 1; Length 662;

Best Local Similarity 100.0%; Pred. No. 2.7e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17

Db 64 TRSRKQTSRLKNIPVH 80

##### RESULT 2

S21382

cell fusion protein - canine distemper virus

C;Species: canine distemper virus

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004

C;Accession: S21382

R;Wild, T.F.; Bernard, A.; Spehner, D.; Villevall, D.; Drillien, R.

submitted to the EMBL Data Library, April 1992

A;Description: Vaccination of mice against canine distemper virus induced encephalitis w

A;Reference number: S21382

A;Accession: S21382

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1-662 <WIL>

A;Cross-references: UNIPROT:P12569; EMBL:X65509; NID:G58853; PIDN:CAA46481.1; PID:G58854

C;Superfamily: parainfluenza virus cell fusion protein

Query Match

100.0%; Score 89; DB 2; Length 662;



Query Match 47.2%; Score 42; DB 2; Length 726;  
Best Local Similarity 52.4%; Pred. No. 41;  
Matches 11; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 1 TRSRK-----QTSHRLKNIP 15  
||| ||| ||| |||  
DB 480 TLRKKLFSNQQRHRTKNIP 500

RESULT 8  
S56936  
vacuolar protein-sorting protein VPS35 - Yeast (*Saccharomyces cerevisiae*)  
N;Alternate names: protein J0580; protein YJUL154C  
C;Species: *Saccharomyces cerevisiae*  
C;Date: 08-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 09-Jul-2004  
C;Accession: S56936; S56937; S31293; S55159; S71643  
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.  
submitted to the Protein Sequence Database, September 1995  
A;Reference number: S56912  
A;Accession: S56936  
A;Molecule type: DNA  
A;Residues: 1-937 <KAT>  
A;Cross-references: UNIPROT:P34110; EMBL:Z49429; MIPS:YJUL154C  
R;Obermaier, B.; Piravandi, E.; Rinke, M.; Domdey, H.  
submitted to the Protein Sequence Database, September 1995  
A;Reference number: S56937  
A;Accession: S56937  
A;Molecule type: DNA  
A;Residues: 829-944 <OBE>  
A;Cross-references: EMBL:Z49429; MIPS:YJUL154C  
R;Paravicini, G.; Horazdovsky, B.F.; Emr, S.D.  
Mol. Biol. Cell 3, 415-427, 1992  
A;Title: Alternative pathways for the sorting of soluble vacuolar proteins in yeast: a v  
A;Reference number: S31293; MUID:92360909; PMID:1498362  
A;Accession: S31293  
A;Molecule type: DNA  
A;Residues: 1-660, 'E', 662-944 <PAR>  
A;Cross-references: EMBL:S42186  
R;Katsoulou, C.; Tzermia, M.; Alexandraki, D.  
submitted to the EMBL Data Library, May 1995  
A;Description: The complete sequence of a 40.7 kb segment located on the left arm of yea  
yeast hypothetical proteins.  
A;Reference number: S55159  
A;Accession: S55159  
A;Molecule type: DNA  
A;Residues: 1-937 <KAW>  
A;Cross-references: EMBL:X87371; NID:g854542; PID:g854543  
R;Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.  
Yeast 12, 787-797, 1996  
A;Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X  
of chromosome XI.  
A;Reference number: S71643; MUID:96408771; PMID:8813765  
A;Accession: S71643  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-937 <KAF>  
A;Cross-references: EMBL:X87371; NID:g854542; PIDN:CRA60801.1; PID:g854543  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995  
C;Genetics:  
A;Gene: SGD:VPS35  
A;Cross-references: SGD:S0003690; MIPS:YJUL154C  
A;Map position: 10L  
C;Function:  
A;Description: required for vacuolar protein sorting  
C;Superfamily: Schizosaccharomycetes probable vacuolar protein sorting-associated protein  
C;Keywords: membrane protein; yeast vacuole

Query Match 47.2%; Score 42; DB 2; Length 944;  
Best Local Similarity 47.1%; Pred. No. 54;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 TRSRKQTSHRLKNIPVH 17

C;Genetics:  
A;Cross-references: SGD:S0001862

A;Map position: 6L  
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YFL032w  
C;Keywords: transmembrane protein  
F;62-78/Domain: transmembrane #status predicted <TM>

Query Match 44.9%; Score 40; DB 2; Length 106;  
Best Local Similarity 53.8%; Pred. No. 12;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 RSRKQTSRLKNIP 14  
DB 16 KSRKQTSRLKNIP 28

RESULT 12  
WMN22

envelope-associated 22K protein - human respiratory syncytial virus

C;Species: human respiratory syncytial virus  
C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C;Accession: B93010; A93009; A04034  
R;Blango, N.; Satake, M.; Venkatesan, S.  
J. Virol. 55, 101-110, 1985

A;Title: mRNA sequence of three respiratory syncytial virus genes encoding two nonstructural proteins  
A;Reference number: A93010; MUID:85237684; PMID:4009789

A;Accession: B93010

A;Molecule type: genomic RNA

A;Residues: 1-194 <ELA>

A;Cross-references: UNIPROT:P04545; GB:M11486; NID:G333925; PIDN:AAB59860.1; PID:G333934

R;Collins, P.L.; Wertz, G.W.

J. Virol. 54, 65-71, 1985  
A;Title: The envelope-associated 22K protein of human respiratory syncytial virus: nucleotide sequence  
A;Reference number: A93009; MUID:85135082; PMID:3838351

A;Accession: A93009

A;Molecule type: genomic RNA

A;Residues: 1-194 <COL>

A;Cross-references: GB:M11486; GB:K01459; GB:K02719; GB:K03348; GB:K03349; GB:M11217; GB:K03349

0.1; PID:G333934

C;Genetics:

A;Gene: 22K

C;Superfamily: respiratory syncytial virus envelope-associated 22K protein

C;Keywords: glycoprotein

F;89,191/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.9%; Score 40; DB 1; Length 194;  
Best Local Similarity 50.0%; Pred. No. 23;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RSRKQTSRLKNIP 15  
DB 140 KNNKQTHLKLRLP 153

RESULT 13  
G82448

sensor histidine kinase/response regulator LuxN VCA0522 [imported] - Vibrio cholerae (strain 569B)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C;Accession: G82448

R;Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Reidberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; et al.

Search completed: September 28, 2005, 17:57:05  
Job time : 19 secs

Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:37:01 ; Search time 91.25 Seconds  
(without alignments)  
72.054 Million cell updates/sec

Title: US-10-705-819B-6

Perfect score: 83

Sequence: 1 SHQYLVTKLPNASLIE 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	100.0	17	3 AAB08081	Aab08081 Amino aci
2	83	100.0	17	8 ADK00487	Adk00487 Immunogen
3	83	100.0	17	8 ADJ84561	Adj84561 T-helper
4	83	100.0	662	3 AAB08102	Aab08102 Amino aci
5	83	100.0	662	4 AAM50124	Aam50124 Canine di
6	76	91.6	662	2 AAR83304	Aar83304 Canine di
7	76	91.6	662	2 AAM47653	Aam47653 Canine di
8	76	91.6	662	7 ADM66120	Adm66120 Vaccinia
9	56	67.5	550	2 AAR34540	Aar34540 F protein
10	56	67.5	550	2 AAR42396	Aar42396 Chicago 1
11	56	67.5	550	2 AAR42394	Aar42394 Moratan h
12	56	67.5	550	2 AAR42397	Aar42397 Consensus
13	56	67.5	550	2 AAR42395	Aar42395 San Diego
14	56	67.5	550	2 AAW94760	Aaw94760 Mutant me
15	56	67.5	550	4 AAB73722	Aab73722 Measles v
16	56	67.5	553	7 ADF86346	Adf86346 Measles v
17	56	67.5	553	7 ADA19559	Ada19559 Measles F
18	52	62.7	15	6 ADA19559	Ada19559 Measles F
19	51	61.4	550	2 AAW94759	Aaw94759 Mutant me
20	49	59.0	17	3 AAB08083	Aab08083 Amino aci
21	49	59.0	17	8 ADK00486	Adk00486 Immunogen
22	49	59.0	17	8 ADJ84560	Adj84560 T-helper
23	47	56.6	17	3 AAB08084	Aab08084 Amino aci
24	47	56.6	17	8 ADK00488	Adk00488 Immunogen
25	47	56.6	17	8 ADJ84562	Adj84562 T-helper

26	47	56.6	22	8 ADK00568	Adk00568 Immunogen
27	47	56.6	22	8 ADK00573	Adk00573 Immunogen
28	47	56.6	23	8 ADK00477	Adk00477 Immunogen
29	47	56.6	23	8 ADK00574	Adk00574 Immunogen
30	47	56.6	26	8 ADK00567	Adk00567 Immunogen
31	47	56.6	26	8 ADJ84716	Adj84716 CTL epito
32	47	56.6	27	8 ADK00473	Adk00473 Immunogen
33	47	56.6	27	8 ADJ84552	Adj84552 CTL epito
34	47	56.6	27	8 ADJ84717	Adj84717 T-helper
35	47	56.6	27	8 ADJ84548	Adj84548 CTL epito
36	47	56.6	27	8 ADJ84719	Adj84719 Lipopepti
37	47	56.6	27	8 ADJ84547	Adj84547 CTL epito
38	47	56.6	46	8 ADK00569	Adk00569 Immunogen
39	47	56.6	47	8 ADK00570	Adk00570 Immunogen
40	44	53.0	178	4 ABG10633	Abg10633 Novel hum
41	44	53.0	604	5 AAU74657	Aau74657 Yeast cyc
42	44	53.0	705	8 ADN19211	Adn19211 Bacterial
43	42	50.6	150	5 ABB49338	Abb49338 Listeria
44	42	50.6	340	4 AAB96780	Aab96780 Putative
45	42	50.6	496	4 ABB61407	Abb61407 Drosophil

## ALIGNMENTS

RESULT 1  
AAB08081  
ID AAB08081 standard; peptide; 17 AA.  
XX AC AAB08081;  
XX AC AAB08081;  
DT 04-DEC-2000 (first entry)  
XX  
DE Amino acid sequence of a helper T cell epitope from CDV.  
XX  
XX T helper cell epitope; CDV; immune response; canine vaccine.  
XX Canine distemper virus.  
XX  
XX WO200046390-A1.  
PN  
PD 10-AUG-2000.  
XX  
PF 07-FEB-2000; 2000WO-AU000070.  
XX  
PR 05-FEB-1999; 99AU-00008533.  
PR 04-AUG-1999; 99AU-00002013.  
XX  
PA (UYME ) UNIV MELBOURNE.  
PA (CSLC-) CSL LTD.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
XX  
PI Jackson DC, Souravi G, Walker J;  
XX WPI; 2000-532904/48.  
DR  
XX Novel T helper cell epitopes derived from canine distemper virus useful for preparation of canine vaccines.  
PT  
XX Claim 1; Page 28; 54pp; English.  
PS  
XX AAB08076-B08101 represent T helper cell epitopes, derived from canine distemper virus (CDV). Compositions comprising these T cell helper epitopes are useful for inducing an immune response in an animal. The epitopes are useful as components of animal, in particular, canine vaccines, either simply as synthetic peptide based vaccines and as additions to vaccines containing more complex antigens  
XX  
SQ Sequence 17 AA;  
Query Match 100.0%; Score 83; DB 3; Length 17;

Best Local Similarity 100.0%; Pred. No. 3.8e-08; Mismatches 0; Indels 0; Gaps 0;  
Matches 17; Conservative 0;

QY 1 SHOYLVIKLIIPNASLIE 17  
|||||  
Db 1 SHOYLVIKLIIPNASLIE 17

RESULT 2  
ADK00487  
ID ADK00487 standard; peptide; 17 AA.

XX AC ADK00487;  
XX DT 06-MAY-2004 (first entry)  
XX DE Immunogenic lipopeptide of the invention #23.  
XX KW T helper cell epitope; B cell epitope; Antibacterial; Antiulcer;  
XX KW Antiinfertility; Vaccine; antibody.  
XX OS Synthetic.  
XX PN WO2004014956-A1.  
XX PD 19-FEB-2004.  
XX PF 12-AUG-2003; 2003WO-AU001018.  
XX PR 12-AUG-2002; 2002US-0402838P.  
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX PI Jackson D, Zeng W;  
XX DR WPI; 2004-238735/22.

XX Novel lipopeptide comprising polypeptide having amino acid sequence of T  
PT helper cell epitope and B cell epitope, conjugated to lipid moieties,  
PT useful for eliciting immune response against group A Streptococcus  
PT antigen.

PS Disclosure; SEQ ID NO 23; 194pp; English.

XX The present invention relates to a lipopeptide comprising polypeptide  
CC conjugated to lipid moieties, where polypeptide contains amino acid  
CC sequence of T helper cell epitope and B cell epitope, where amino acid  
CC sequences are different, and internal lysine residues or internal lysine  
CC analog residues for covalent attachment of each of lipid moieties through  
CC &agr;g; amino group or terminal side chain group of lysine or lysine  
CC analog. The peptides are useful in eliciting the production of antibody  
CC against an antigenic B cell epitope in a subject, and are useful for  
CC antibody production, synthetic vaccine production, diagnostic method  
CC employing antibodies and antibody ligands and immunotherapy for  
CC veterinary and human medicine. The method efficiently elicits the  
CC production of antibody against antigenic B cell epitope. The present  
CC sequence represents a novel immunogenic lipopeptide comprising T helper  
CC and B cell epitopes.

SQ Sequence 17 AA;

Query Match 100.0%; Score 83; DB 8; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOYLVIKLIIPNASLIE 17  
|||||  
Db 1 SHOYLVIKLIIPNASLIE 17

RESULT 3  
ADJ84561  
ID ADJ84561 standard; peptide; 17 AA.

XX AC ADJ84561;  
XX DT 06-MAY-2004 (first entry)  
XX DE T-helper epitope peptide SEQ ID NO:19.  
XX KW lipopeptide; lipid moiety; T helper cell epitope; Th epitope;  
XX KW cytotoxic T cell epitope; CTL epitope; immune response; immunisation;  
XX KW influenza; hepatitis C virus; Listeria monocytogenes; vaccine; cancer;  
XX KW cytostatic; virucide.  
XX OS Canine distemper virus.  
XX OS Synthetic.  
XX PN WO2004014957-A1.  
XX PD 19-FEB-2004.  
XX PF 12-AUG-2003; 2003WO-AU001019.  
XX PR 12-AUG-2002; 2002US-0403328P.  
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX PI Jackson D, Zeng W;  
XX DR WPI; 2004-238736/22.

XX New lipopeptide comprising a polypeptide comprising an amino acid  
PT sequence of a T helper cell and cytotoxic T cell epitope, useful for  
PT preparing a composition for treating or preventing cancer, or hepatitis C  
PT virus or influenza virus.

PS Disclosure; SEQ ID NO 19; 166pp; English.

XX The present invention describes a lipopeptide comprising a polypeptide  
CC conjugated to one or more lipid moieties, where the polypeptide comprises  
CC an amino acid sequence that comprises: (a) the amino acid sequence of a T  
CC helper cell (Th) epitope and the amino acid sequence of a cytotoxic T  
CC cell (CTL) epitope, where the amino acid sequences are different; and (b)  
CC one or more internal lysine residues or internal lysine analogue residues  
CC for covalent attachment of each of the lipid moieties via the epsilon-  
CC amino group or terminal side-chain group of the lysine or lysine  
CC analogue. Each of the one or more lipid moieties is covalently attached  
CC to an epsilon-amino group of the one or more internal lysine residues or  
CC to a terminal side-chain group of the one or more internal lysine  
CC analogue residues. Also described: (1) producing a lipopeptides; (2) a  
CC composition comprising the lipopeptide and an excipient or diluent; (3)  
CC eliciting an immune response in a subject; (4) immunising a subject  
CC against influenza or hepatitis C virus or Listeria monocytogenes; (5) a  
CC vaccine against an influenza virus or hepatitis C virus or Listeria  
CC monocytogenes comprising the lipopeptides, where the CTL epitope is from  
CC an influenza virus or hepatitis C virus or Listeria monocytogenes protein  
CC; (6) treating or preventing cancer; and (7) a vaccine against cancer  
CC comprising the lipopeptide, where the CTL epitope is a tumour-specific  
CC CTL epitope. The lipopeptide has cytostatic and virucide activities, and  
CC can be used in vaccine. The lipopeptide is useful for preparing a  
CC composition for treating or preventing cancer, or hepatitis C virus or  
CC influenza virus. The present sequence is used in the exemplification of  
CC the present invention.

SQ Sequence 17 AA;

Query Match 100.0%; Score 83; DB 8; Length 17;  
Best Local Similarity 100.0%; Pred. No. 3.8e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHOYLVIKLIIPNASLIE 17  
|||||  
Db 1 SHOYLVIKLIIPNASLIE 17



```

RESULT 4
AAB08102
ID AAB08102 standard; protein; 662 AA.
XX
XX AC AAB08102;
XX
XX DT 04-DEC-2000 (first entry)
XX
XX DE Amino acid sequence of a fusion protein of canine distemper virus.
XX
XX KW T helper cell epitope; CDV; immune response; canine vaccine.
XX
XX OS Canine distemper virus.
XX
XX PN WO200046390-A1.
XX
XX PD 10-AUG-2000.
XX
XX PF 07-FEB-2000; 2000WO-AU000070.
XX
XX PR 05-FEB-1999; 39AU-00008533.
XX
XX PR 04-AUG-1999; 99AU-00002013.
XX
XX PA (UYME ) UNIV MELBOURNE.
XX
XX PA (CSLC-) CSL LTD.
XX
XX PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
XX PI Jackson DC, Souravi G, Walker J;
XX
XX WPI; 2000-532904/48.
XX
XX PT Novel T helper cell epitopes derived from canine distemper virus useful
XX for preparation of canine vaccines.
XX
XX PS Example 1; Fig 1; 54pp; English.
XX
XX CC The present sequence represents a fusion protein of canine distemper
XX virus (CDV). The protein was used to identify peptides AAB08076-B08101,
XX which are T helper cell epitopes. Compositions comprising these T cell
XX helper epitopes are useful for inducing an immune response in an animal.
XX The epitopes are useful as components of animal, in particular, canine
XX vaccines, either simply as synthetic peptide based vaccines and as
XX additions to vaccines containing more complex antigens
XX
XX SQ Sequence 662 AA;

Query Match 100.0%; Score 83; DB 3; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17
|||||
DB 162 SHQYLVKILPNASLIE 178

RESULT 5
AAM50124
ID AAM50124 standard; protein; 662 AA.
XX
XX AC AAM50124;
XX
XX DT 21-DEC-2001 (first entry)
XX
XX DE Canine distemper virus recombinant fusion protein PCDFV662.
XX
XX KW CDV; fusion protein; PCDFV662; antigen; immune status;
XX vaccination status; dog.
XX
XX OS Canine distemper virus.
XX
XX PN WO200166568-A2.

Query Match 100.0%; Score 83; DB 3; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHQYLVKILPNASLIE 17
|||||
DB 162 SHQYLVKILPNASLIE 178

RESULT 6
AAR83304
ID AAR83304 standard; protein; 662 AA.
XX
XX AC AAR83304;
XX
XX DT 25-MAR-2003 (revised)
XX
XX DT 29-MAY-1996 (first entry)
XX
XX DE Canine distemper virus fusion glycoprotein.
XX
XX KW Avian influenza virus; rabies; canine distemper virus; CDV; measles;
XX antibody; vaccine; therapy; TROVAC.
XX
XX OS Synthetic.
XX
XX PN WO9527780-A1.
XX
XX PD 19-OCT-1995.
XX
XX PF 06-APR-1995; 95WO-US004394.
XX
XX PR 06-APR-1994; 94US-00224657.
XX
XX PR 05-APR-1995; 95US-00416646.
XX
XX PA (VIRO-) VIROGENETICS CORP.
XX
XX PI Paoletti E, Tartaglia J, Taylor J, Gettig R;

```

13-SEP-2001.  
 07-MAR-2001; 2001WO-US007251.  
 09-MAR-2000; 2000US-00521738.  
 (HESK-) HESKA CORP.  
 (COLS ) UNIV COLORADO STATE RES FOUND.  
 Jensen WA, Lappin MR, Rosen DK, Andrews JS;  
 WPI; 2001-639000/73.  
 N-PSDB; AAH27071.  
 Determining immune status or vaccination status of an animal to e.g.  
 calicivirus comprises using a recombinant viral antigen.  
 Claim 4; Page 129-131; 132pp; English.  
 The present sequence is that of canine distemper virus fusion protein  
 recombinant antigen, PCDFV662. The recombinant antigen, fused to an N-  
 terminal His tag, was produced in Escherichia coli cells transformed by a  
 recombinant vector comprising nucleic acid nCDVFI986 (see AAH27071).  
 PCDFV662 is an example of a recombinant infectious agent antigen that can  
 be used in the method of the invention to determine the immune status of  
 an animal. The method involves contacting a biological specimen of an  
 animal (cat, dog or horse) with a recombinant antigen, and detecting the  
 presence or absence of a complex between the recombinant antigen and an  
 antibody present in the sample. The method determines whether the animal  
 is protected against disease or should be vaccinated. Recombinant  
 antigens (see AAM50107-24), nucleic acids encoding them (see AAH27054-  
 71), methods of producing them, and assay methods are provided  
 Sequence 662 AA;

XX WPI; 1995-366385/47.  
 DR N-PSDB; AAT00521.  
 XX  
 PT New modified recombinant viruses - contg. a canine distemper virus  
 PT antigen or measles virus M or N antigen in a non-essential region.  
 XX  
 PS Example 15; Fig 15; 194pp; English.  
 XX  
 CC This sequence represents the H6 promoted canine distemper virus (CDV)  
 CC fusion glycoprotein sequence. The DNA encoding this sequence is contained  
 CC in the plasmid pSCDVA. The encoding sequence was removed from this  
 CC plasmid and inserted into the C5 locus of another plasmid. The vector can  
 CC then be used to introduce the exogenous DNA into a fowlpox virus. This  
 CC procedure can be performed using exogenous DNA encoding a measles virus M  
 CC or N antigen. The recombinant viruses can then be used to induce an  
 CC antigenic or immunological response. The expression products of these  
 CC recombinants, and the antibodies produced can be used in binding assays  
 CC to determine the presence of CDV or measles virus in a sample. The  
 CC attenuated virulence of the viruses reduces the possibility of a runaway  
 CC infection due to the vaccination, in a vaccinated individual. The viruses  
 CC also reduce the transmission from vaccinated to unvaccinated individuals,  
 CC and reduces environmental contamination. (Updated on 25-MAR-2003 to  
 CC correct PR field.)  
 XX  
 CC Sequence 662 AA;  
 SQ

Query Match 91.6%; Score 76; DB 2; Length 662;  
 Best Local Similarity 88.2%; Pred. No. 5.3e-05;  
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SHQYLVKILPNASLIE 17  
 DB 162 SHQYLVKILMPNVSLIE 178  
 |||||:|||||  
 |||||:|||||

RESULT 7  
 AAM47653  
 ID AAM47653 standard; protein; 662 AA.  
 XX  
 AC AAM47653;  
 XX  
 DT 21-FEB-2002 (first entry)  
 DE Canine distemper virus, CDV, F.  
 XX  
 KW Virucide; vaccine; virus; virulence; canine distemper virus; CDV;  
 KW measles; dog.  
 XX  
 OS Canine distemper virus.  
 XX  
 FN US6309647-B1.  
 XX  
 PD 30-OCT-2001.  
 XX  
 PF 15-JUL-1999; 99US-00354138.  
 XX  
 PR 15-JUL-1999; 99US-00354138.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Paoletti E, Tartaglia J, Taylor J, Gettig R;  
 XX  
 DR WPI; 2002-040232/05.  
 DR N-PSDB; AB198922.  
 XX  
 PT Novel virus, useful for inducing immune response in dog against CDV,  
 PT comprises the modified recombinant virus having attenuated virulence  
 PT comprising exogenous DNA sequences encoding antigens of canine distemper  
 PT virus (CDV) or measles virus.  
 XX  
 PS Example 15; Fig 15; 147pp; English.  
 XX

CC The present invention relates to modified recombinant viruses, comprising  
 CC inactivated virus-encoded genetic functions so that the viruses have  
 CC attenuated virulence, yet retained efficiency. The viruses can contain  
 CC DNA encoding a canine distemper virus (CDV) antigen or measles M or N  
 CC antigen. The recombinant viruses are useful for inducing an antigenic or  
 CC immunological response in a dog or other carnivore against CDV. The  
 CC present sequence was used in an example from the present invention  
 XX  
 SQ Sequence 662 AA;  
 Query Match 91.6%; Score 76; DB 5; Length 662;  
 Best Local Similarity 88.2%; Pred. No. 5.3e-05;  
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 SHQYLVKILPNASLIE 17  
 DB 162 SHQYLVKILMPNVSLIE 178  
 |||||:|||||  
 |||||:|||||

RESULT 8  
 ADM66120  
 ID ADM66120 standard; protein; 662 AA.  
 XX  
 AC ADM66120;  
 XX  
 DT 03-JUN-2004 (first entry)  
 DE Vaccinia virus H6-promoted CDV Fusion (F) protein #1.  
 XX  
 KW Canine distemper virus; NYVAC; TROVAC; ALVAC; rabies glycoprotein G;  
 KW rabies; vaccine.  
 XX  
 OS Canine distemper virus; Onderstepoort strain.  
 XX  
 FN US2003082204-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 13-SEP-2001; 2001US-00951061.  
 XX  
 PR 20-NOV-1990; 90US-00621614.  
 PR 07-MAR-1991; 91US-00666056.  
 PR 11-JUN-1991; 91US-00713967.  
 PR 22-OCT-1991; 91US-00776867.  
 PR 06-MAR-1992; 92US-00847951.  
 PR 31-AUG-1992; 92US-00938283.  
 PR 08-JUN-1993; 93US-00073962.  
 PR 12-AUG-1993; 93US-00105483.  
 PR 06-APR-1994; 94US-00224657.  
 PR 15-JUL-1999; 99US-00354138.  
 XX  
 PA (AVET ) AVENTIS PASTEUR.  
 XX  
 PI Paoletti E, Tartaglia J, Taylor J, Gettig R;  
 XX  
 DR WPI; 2003-567445/53.  
 DR N-PSDB; ADM66154.  
 XX  
 PT New recombinant viruses comprising exogenous DNA encoding rabies  
 PT glycoprotein G useful for eliciting protective immunity against rabies  
 PT virus in a carnivore.  
 XX  
 PS Example 15; SEQ ID NO 86; 93pp; English.  
 XX  
 CC The invention relates to recombinant vaccinia and canarypox viruses  
 CC comprising exogenous DNA encoding rabies glycoprotein G in a nonessential  
 CC region of the virus genome. Also included are a recombinant vaccinia virus  
 CC (comprising exogenous DNA encoding rabies glycoprotein G in a  
 CC nonessential region of the virus genome, where at least one open reading  
 CC frame (ORF) selected from J2, B13, +B14R, A26L, 156R, C7L-K1L, and 14L is  
 CC deleted from the virus), a recombinant canarypox virus (produced by  
 CC attenuation through multiple serial passages on chick embryo fibroblasts,  
 CC subjecting a master seed from to successive plaque purifications under

agar and amplifying a plaque clone through multiple additional passages, where the virus contains exogenous DNA encoding rabies glycoprotein G in a nonessential region of the virus genome), inducing an antigenic or immunological response in a carnivore against rabies virus (by administering to the dog, cat or other carnivore a composition comprising the virus above in a mixture with a carrier), expressing a gene product in a cell cultured in vitro by introducing into the cell a virus of the invention. One or more (optionally all) ORFs selected from a thymidine kinase gene, a haemorrhagic region, an A type inclusion body, a haemagglutinin gene, a host range region, and a ribonucleotide reductase large subunit gene, may also be deleted. The attenuated Vaccinia virus is termed a NYVAC virus. The attenuated canarypox virus is termed ALVAC recombinant virus. Also included are attenuated fowl pox viruses termed TROVAC. The recombinant viruses are useful as vaccines for protecting a dog, cat or other carnivore against rabies. The modified recombinant viruses are effective as vaccines and are safer than some other recombinant viruses due to the deletion of genes affecting virulence that are not essential for virus growth in tissue culture. The present sequence is a Canine distemper virus protein expressed from an attenuated virus of the invention.

CC Sequence 662 AA;

Query Match 91.6%; Score 76; DB 7; Length 662;  
Best Local Similarity 88.2%; Pred. No. 5.3e-05;  
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 SHQYLVKILIPNASLIE 17  
Db 162 SHQYLVKILIPNASLIE 178

RESULT 9  
AAR34540  
ID AAR34540 standard; protein; 550 AA.

XX AC AAR34540;

XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 19-AUG-1993 (first entry)

XX DE F protein of attenuated measles virus strain AIK-C.

XX KW paramyxoviridae; RNA virus; attenuation; vaccine.

XX OS Measles virus.

XX PN EP540135-A2.

XX PD 05-MAY-1993.

XX PF 10-MAR-1992; 92EP-00302004.

XX PR 14-OCT-1991; 91JP-00293625.

XX PA (KITA) KITASATO INST.

XX PI Sasaki K, Mori T, Makino S;

XX DR WPI; 1993-145503/18.

XX DR N-PSDB; AAR34540.

XX PT New attenuated measles vaccine virus strain - retains high immunogenicity with reduced pyrogenicity and having no neurological complications.

XX PS Disclosure; Page 14-18; 47pp; English.

XX CC The sequence of the entire genome of the attenuated measles virus strain AIK-C has been determined and contains 6 open reading frames. The sequence of the F protein was deduced from the fourth ORF. The virus retains high immunogenicity with reduced pyrogenicity and no neurological complications. See AAR34537-R34541 and AAR39592. (Updated on 25-MAR-2003

CC to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)  
XX SQ Sequence 550 AA;

Query Match 67.5%; Score 56; DB 2; Length 550;  
Best Local Similarity 68.8%; Pred. No. 0.22;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQYLVKILIPNASLI 16  
Db 50 SHQSLVILKIPNITLL 65

RESULT 10

AAR42396

ID AAR42396 standard; protein; 550 AA.

XX AC AAR42396;

XX DT 24-OCT-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 13-MAY-1994 (first entry)

XX DS Chicago 1 haemagglutinin fusion protein.

XX KW Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;

XX KM vaccine; infection; consensus polypeptide.

XX OS Measles virus; strain Chicago 1 f.

XX PN W09321325-A1.

XX PD 28-OCT-1993.

XX PF 08-APR-1993; 93WO-US003209.

XX PR 08-APR-1992; 92US-00866033.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX PI Rota JS, Bellini WJ;

XX DR WPI; 1993-351735/44.

XX DR N-PSDB; AAR51097.

XX PT Haemagglutinin and fusion glycoprotein of several wild-type measles strains - used to construct vaccines for measles infection.

XX PS Disclosure; Page 68-71; 119pp; English.

XX CC HA (AAR51088-94) and fusion glycoprotein (AAR51095-97) sequences of several wild-type measles strains are given. Shared amino acid variations in wild-type measles glycoproteins are identified in five wild-type measles viruses. A consensus polypeptide, the amino acid sequence of which reflects variation common to more than one wild-type strain, is the basis for constructing live attenuated vaccines, or recombinant vaccines to replace older, less efficacious vaccines. Immunological reagents useful in differentiating wild-type measles strains from other CC known strains can also be produced. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS field)

XX SQ Sequence 550 AA;

Query Match 67.5%; Score 56; DB 2; Length 550;  
Best Local Similarity 68.8%; Pred. No. 0.22;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SHQYLVKILIPNASLI 16  
Db 50 SHQSLVILKIPNITLL 65

RESULT 11

```

AAR42394
ID AAR42394 standard; protein; 550 AA.
XX
AC AAR42394;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 13-MAY-1994 (first entry)
XX
DE Moraten haemagglutinin fusion protein.
XX
KW Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;
KW vaccine; infection; consensus polypeptide.
XX
OS Measles virus strain Moraten; f.
XX
PN WO9321325-A1.
XX
PD 28-OCT-1993.
XX
PF 08-APR-1993; 93WO-US003209.
XX
PR 08-APR-1992; 92US-00866033.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Rota JS, Bellini WJ;
PI PI
PI WPI; 1993-351735/44.
XX
DR Haemagglutinin and fusion glycoprotein of several wild-type measles
XX strains - used to construct vaccines for measles infection.
XX
PS Claim 13; Page 78-80; 119pp; English.
XX
CC HA (AAO51088-94) and fusion glycoprotein (AAO51095-97) sequences of
XX several wild-type measles strains are given. Shared amino acid variations
XX in wild-type measles glycoproteins are identified in five wild-type
XX measles viruses. A consensus polypeptide, the amino acid sequence of
XX which reflects variation common to more than one wild-type strain, is
XX the basis for constructing live attenuated vaccines, or recombinant
XX vaccines to replace older, less efficacious vaccines. Immunological
XX reagents useful in differentiating wild-type measles strains from other
XX known strains can also be produced. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
SQ Sequence 550 AA;
XX
Query Match 67.5%; Score 56; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.22;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQYLVIKLPNASLI 16
DB 50 SHQSLVILKMPNITLL 65

RESULT 13
AAR42395
ID AAR42395 standard; protein; 550 AA.
XX
AC AAR42395;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 13-MAY-1994 (first entry)
XX
DE San Diego haemagglutinin fusion protein.
XX
KW Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;
KW vaccine; infection; consensus polypeptide.
XX
OS Measles virus; strain San Diego f.
XX
PN WO9321325-A1.
XX
PD 28-OCT-1993.
XX
PF 08-APR-1993; 93WO-US003209.
XX
PR 08-APR-1992; 92US-00866033.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Rota JS, Bellini WJ;
PI PI
PI WPI; 1993-351735/44.
XX
DR Haemagglutinin and fusion glycoprotein of several wild-type measles
XX strains - used to construct vaccines for measles infection.
XX
PS Disclosure; Page 58-61; 119pp; English.
XX
CC HA (AAO51088-94) and fusion glycoprotein (AAO51095-97) sequences of
XX several wild-type measles strains are given. Shared amino acid variations
XX in wild-type measles glycoproteins are identified in five wild-type
XX measles viruses. A consensus polypeptide, the amino acid sequence of
XX which reflects variation common to more than one wild-type strain, is
XX the basis for constructing live attenuated vaccines, or recombinant
XX vaccines to replace older, less efficacious vaccines. Immunological
XX reagents useful in differentiating wild-type measles strains from other
XX known strains can also be produced. (Updated on 25-MAR-2003 to correct PN
XX field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 550 AA;
XX
Query Match 67.5%; Score 56; DB 2; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.22;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQYLVIKLPNASLI 16
DB 50 SHQSLVILKMPNITLL 65

RESULT 12
AAR42397
ID AAR42397 standard; protein; 550 AA.
XX
AC AAR42397;
XX
DT 25-MAR-2003 (revised)
DT 13-MAY-1994 (first entry)
XX
DE Consensus haemagglutinin fusion protein.
XX
KW Haemagglutinin; HA; fusion glycoprotein; wild-type; measles virus;
KW vaccine; infection; consensus polypeptide.
XX
OS Measles virus.

```

XX Haemagglutinin and fusion glycoprotein of several wild-type measles  
 PT strains - used to construct vaccines for measles infection.  
 XX  
 PS Disclosure; Page 63-66; 119pp; English.  
 XX  
 CC HA (AAQ51088-94) and fusion glycoprotein (AAQ51095-97) sequences of  
 CC several wild-type measles strains are given. Shared amino acid variations  
 CC in wild-type measles glycoproteins are identified in five wild-type  
 CC measles viruses. A consensus polypeptide, the amino acid sequence of  
 CC which reflects variation common to more than one wild-type strain, is  
 CC the basis for constructing live attenuated vaccines, or recombinant  
 CC vaccines to replace older, less efficacious vaccines. Immunological  
 CC reagents useful in differentiating wild-type measles strains from other  
 CC known strains can also be produced. (Updated on 25-MAR-2003 to correct PN  
 CC field.) (Updated on 24-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 550 AA;

Query Match 67.5%; Score 56; DB 2; Length 550;  
 Best Local Similarity 68.8%; Pred. No. 0.22;  
 Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16  
 ||| |||||:|:|:  
 Db 50 SHQSLVTKLMPNITLL 65

RESULT 14  
 AAW94760  
 ID AAW94760 standard; protein; 550 AA.  
 XX  
 AC AAW94760;

DT 20-APR-1999 (first entry)

DE Mutant measles virus F protein antigen (strain NA).

XX Measles; H protein; F protein; antigen; mutant; vaccine; detection;  
 KW diagnostic reagent; attenuated.

XX Measles virus.

XX Key Location/Qualifiers  
 FT Misc-difference 362  
 FT /note= "encoded by TCC"

XX WO9855627-A1.

XX 10-DEC-1998.

XX 04-JUN-1998; 98WO-JP002481.

XX 04-JUN-1997; 97JP-00184285.

XX (OSAU) UNIV OSAKA.

XX Ueda S, Watanabe M, Kawanishi H;

XX WPI; 1999-070272/06.

XX N-PSDB; AAX05591.

XX Mutant measles virus H and F protein antigens and genes encoding them -  
 PT for production of attenuated virus or genetic vaccines and of diagnostic  
 PT reagents.

XX Claim 1; Page 80-83; 93pp; Japanese.

XX This represents a mutant measles virus F protein antigen. This protein is  
 CC isolated from the NA (epidemic) strain of measles virus. The invention  
 CC provides mutant measles virus antigenic proteins such as the H protein  
 CC from the CAM-70 (attenuated) or NA (epidemic) strain of measles virus and  
 CC the F protein from the CAM-70 or NA strain and nucleic acid sequences

CC encoding the proteins. These antigenic proteins are used in the  
 CC production of vaccines for measles and the production of diagnostic  
 CC reagents (e.g. measles antigens for the detection of specific antibodies  
 CC in the blood of patients). The vaccines may be recombinant or other live  
 CC attenuated measles virus strains or may be genetic vaccines (e.g.  
 CC adenovirus vectors)

XX Sequence 550 AA;

Query Match 67.5%; Score 56; DB 2; Length 550;  
 Best Local Similarity 68.8%; Pred. No. 0.22;

Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SHQYLVIKLIPNASLI 16  
 ||| |||||:|:|:  
 Db 50 SHQSLVTKLMPNITLL 65

RESULT 15

AAAB73722

ID AAB73722 standard; protein; 550 AA.

XX AAB73722;

DT 11-SEP-2001 (first entry)

DE Measles virus strain AIK-C F protein.

XX F protein; fusion protein; cell fusion ability; attenuated strain;  
 KW morbillivirus; vaccine; mutant; mutein.

XX Measles virus strain AIK-C.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 278

FT /note= "This residue is critical for the cell fusing  
 FT activity of the F protein, and is substituted by Phe in  
 FT the F protein of the virulent Edmonston strain"

XX WO200138537-A1.

XX 31-MAY-2001.

XX 18-OCT-2000; 2000WO-JP007233.

XX 19-NOV-1999; 99JP-00330099.

XX (KITA) KITASATO INST.

XX Komase K, Nakayama T, Aizawa C;

XX WPI; 2001-367692/38.

XX N-PSDB; AAH24296.

XX DNA for constructing viruses with low cell fusing ability useful as  
 PT vaccines.

XX Example 1; Page 34-36; 41pp; Japanese.

XX The invention relates to the construction of attenuated morbillivirus  
 CC strains which have low cell fusing ability. The morbilliviruses include  
 CC pathogenic animal viruses such as measles virus, canine distemper virus  
 CC (CDV), phocine distemper virus (PDV), and rinderpest virus (RPV). The  
 CC protein responsible for cell fusing is the F (fusion) protein, in which  
 CC the amino acid corresponding to residue 278 of the measles virus is  
 CC critical for normal cell fusing activity. Accordingly, the invention  
 CC specifically claims a measles virus F protein which contains any amino  
 CC acid except Phe at residue 278, and proteins with at least 65% homology  
 CC to such F proteins. The invention also encompasses DNA encoding the F  
 CC proteins of the invention and vectors and morbilliviruses comprising such  
 CC DNA. Medical compositions comprising an attenuated morbillivirus of the F  
 CC invention are useful as vaccines. The present sequence represents the F

CC protein from the attenuated AIK-C strain of measles virus  
XX  
SQ Sequence 550 AA;

Query Match 67.5%; Score 56; DB 4; Length 550;  
Best Local Similarity 68.8%; Pred. No. 0.22;  
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SHQYLVIKLIPNASLI 16  
||| ||||| :|| :||  
Db 50 SHQSLVIKLMFNITLL 65

Search completed: September 28, 2005, 17:50:24  
Job time : 92.25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:37:01 ; Search time 91.25 Seconds  
(without alignments)  
72.054 Million cell updates/sec

Title: US-10-705-819B-5  
Perfect score: 89  
Sequence: 1 TRSRKQTSRLKNIPVH 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseqp16Dec04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	17	3 AAB08080	AAB08080 Amino aci
2	89	100.0	17	8 ADK00506	ADK00506 Immunogen
3	89	100.0	17	8 ADJ84580	ADJ84580 T-helper
4	89	100.0	662	2 AAR83304	AAR83304 Canine di
5	89	100.0	662	3 AAB08102	AAB08102 Amino aci
6	89	100.0	662	4 AAM50124	AAM50124 Canine di
7	89	100.0	662	5 AAM47653	AAM47653 Canine di
8	89	100.0	662	7 ADM66120	ADM66120 Vaccinia
9	44.5	50.0	280	4 ABG00101	ABG00101 Novel hum
10	44.5	50.0	280	4 ABG00285	ABG00285 Novel hum
11	44	49.4	103	5 ABP40732	ABP40732 Staphyloc
12	44	49.4	103	8 ADS08038	ADS08038 Staphyloc
13	43	48.3	209	8 ADS27523	ADS27523 Bacterial
14	42	47.2	186	2 AAR24190	AAR24190 Bovine RS
15	42	47.2	354	8 ADS21233	ADS21233 Bacterial
16	42	47.2	726	6 ABU28584	ABU28584 Protein e
17	42	47.2	944	7 ADK64818	ADK64818 Disease t
18	41	46.1	186	5 ABG97497	ABG97497 Human NOV
19	41	46.1	186	5 AAO21665	AAO21665 Human sec
20	41	46.1	188	5 ABG97498	ABG97498 Human NOV
21	41	46.1	194	4 ADG27804	ADG27804 Human nov
22	41	46.1	403	6 AAE30053	AAE30053 Arabidops
23	41	46.1	403	7 ADC46717	ADC46717 Thalecres
24	41	46.1	403	8 ADI43593	ADI43593 Plant tra
25	41	46.1	403	8 ADO02069	ADO02069 Thalecres

26	41	46.1	470	7 ADE80781	Ade80781 Microsate
27	41	46.1	471	7 ADE80782	Ade80782 Microsate
28	41	46.1	501	7 ADE80780	Ade80780 Microsate
29	41	46.1	630	4 AAB93459	AAB93459 Human pro
30	41	46.1	630	7 ADE80779	Ade80779 Microsate
31	41	46.1	1294	4 AAM41110	AAM41110 Human pol
32	41	46.1	1323	7 ADB82781	ADB82781 Human pro
33	41	46.1	1700	4 AAM39324	AAM39324 Human pol
34	41	46.1	1818	7 ADJ69293	ADJ69293 Human hea
35	40	44.9	60	4 AAU14779	AAU14779 Novel bon
36	40	44.9	74	4 AAU65519	AAU65519 Propionib
37	40	44.9	74	6 ABM62038	ABM62038 Propionib
38	40	44.9	92	5 ABP01229	ABP01229 Human ORF
39	40	44.9	96	4 AAU41556	AAU41556 Propionib
40	40	44.9	96	6 ABM38075	ABM38075 Propionib
41	40	44.9	188	2 AAY38708	AAY38708 Neisseria
42	40	44.9	188	2 AAY38710	AAY38710 Neisseria
43	40	44.9	194	1 AAP70785	AAP70785 Sequence
44	40	44.9	194	1 AAP70477	AAP70477 Sequence
45	40	44.9	194	2 AAR25303	AAR25303 HRSV 22K

## ALIGNMENTS

RESULT 1  
AAB08080  
ID AAB08080 standard; peptide; 17 AA.  
XX AAB08080;  
AC AAB08080;  
XX  
DT 04-DEC-2000 (first entry)  
XX  
DE Amino acid sequence of a helper T cell epitope from CDV.  
XX  
KW T helper cell epitope; CDV; immune response; canine vaccine.  
XX  
OS Canine distemper virus.  
XX  
FN WO200046390-A1.  
PN 10-AUG-2000.  
PD  
XX  
PF 07-FEB-2000; 2000WO-AU0000070.  
XX  
PR 05-FEB-1999; 99AU-00008533.  
PR 04-AUG-1999; 99AU-00002013.  
XX  
PA (UPME ) UNIV MELBOURNE.  
PA (CSLC-) CSL LTD.  
PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.  
PI Jackson DC, Souravi G, Walker J;  
XX WPI; 2000-532904/48.  
DR  
XX Novel T helper cell epitopes derived from canine distemper virus useful  
PT for preparation of canine vaccines.  
XX  
PS Claim 1; Page 28; 54pp; English.  
XX  
CC AAB08076-B08101 represent T helper cell epitopes, derived from canine  
CC distemper virus (CDV). Compositions comprising these T cell helper  
CC epitopes are useful for inducing an immune response in an animal. The  
CC epitopes are useful as components of animal, in particular, canine  
CC vaccines, either simply as synthetic peptide based vaccines and as  
CC additions to vaccines containing more complex antigens  
XX  
SQ Sequence 17 AA;  
Query Match 100.0%; Score 89; DB 3; Length 17;

```

Best Local Similarity 100.0%; Pred. No. 9.4e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||
Db 1 TRSRKQTSRLKNIPVH 17

RESULT 2
ID ADK00506
XX ADK00506 standard; peptide; 17 AA.
AC ADK00506;
XX
XX 06-MAY-2004 (first entry)
XX
XX Immunogenic lipopeptide of the invention #42.
XX
XX T helper cell epitope; B cell epitope; Antibacterial; Antitumor;
XX Antiinfectivity; Vaccine; antibody.
XX
XX Synthetic.
XX
XX WO2004014956-A1.
XX
XX 19-FEB-2004.
XX
XX 12-AUG-2003; 2003WO-AU001018.
XX
XX 12-AUG-2002; 2002US-0402838P.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
XX Jackson D, Zeng W;
XX
XX WPI; 2004-238735/22.
XX
XX Novel lipopeptide comprising polypeptide having amino acid sequence of T
XX helper cell epitope and B cell epitope, conjugated to lipid moieties,
XX useful for eliciting immune response against group A Streptococcus
XX antigen.
XX
XX Disclosure; SEQ ID NO 42; 194pp; English.
XX
XX The present invention relates to a lipopeptide comprising polypeptide
XX conjugated to lipid moieties, where polypeptide contains amino acid
XX sequence of T helper cell epitope and B cell epitope, where amino acid
XX sequences are different, and internal lysine residues or internal lysine
XX analog residues for covalent attachment of each of lipid moieties through
XX asparagine or terminal side chain group of lysine or lysine
XX analog. The peptides are useful in eliciting the production of antibody
XX against an antigenic B cell epitope in a subject, and are useful for
XX antibody production, synthetic vaccine production, diagnostic method
XX employing antibodies and antibody ligands and immunotherapy for
XX veterinary and human medicine. The method efficiently elicits the
XX production of antibody against antigenic B cell epitope. The present
XX sequence represents a novel immunogenic lipopeptide comprising T helper
XX and B cell epitopes.
XX
XX Sequence 17 AA;
XX
XX Query Match 100.0%; Score 89; DB 8; Length 17;
XX Best Local Similarity 100.0%; Pred. No. 9.4e-09;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||
Db 1 TRSRKQTSRLKNIPVH 17

RESULT 3
ADJ84580
ID ADJ84580 standard; peptide; 17 AA.

```

```

XX AC ADJ84580;
XX DT 06-MAY-2004 (first entry)
XX DE T-helper epitope peptide SEQ ID NO:38.
XX KW lipopeptide; lipid moiety; T helper cell epitope; Th epitope;
XX cytotoxic T cell epitope; CTL epitope; immune response; immunisation;
XX influenza; hepatitis C virus; Listeria monocytogenes; vaccine; cancer;
XX cytostatic; virucide.
XX OS Canine distemper virus.
XX OS Synthetic.
XX PN WO2004014957-A1.
XX PD 19-FEB-2004.
XX PF 12-AUG-2003; 2003WO-AU001019.
XX PR 12-AUG-2002; 2002US-0403328P.
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX PI Jackson D, Zeng W;
XX DR WPI; 2004-238736/22.
XX
XX New lipopeptide comprising a polypeptide comprising an amino acid
XX sequence of a T helper cell and cytotoxic T cell epitope, useful for
XX preparing a composition for treating or preventing cancer, or hepatitis C
XX virus or influenza virus.
XX
XX Disclosure; SEQ ID NO 38; 166pp; English.
XX
XX The present invention describes a lipopeptide comprising a polypeptide
XX conjugated to one or more lipid moieties, where the polypeptide comprises
XX an amino acid sequence that comprises: (a) the amino acid sequence of a T
XX helper cell (Th) epitope and the amino acid sequence of a cytotoxic T
XX cell (CTL) epitope, where the amino acid sequences are different; and (b)
XX one or more internal lysine residues or internal lysine analogue residues
XX for covalent attachment of each of the lipid moieties via the epsilon-
XX amino group or terminal side-chain group of the lysine or lysine
XX analogue. Each of the one or more lipid moieties is covalently attached
XX to an epsilon-amino group of the one or more internal lysine residues or
XX to a terminal side-chain group of the one or more internal lysine
XX analogue residues. Also described: (1) producing a lipopeptide; (2) a
XX composition comprising the lipopeptide and an excipient or diluent; (3)
XX eliciting an immune response in a subject; (4) immunising a subject
XX against influenza or hepatitis C virus or Listeria monocytogenes; (5) a
XX vaccine against an influenza virus or hepatitis C virus or Listeria
XX monocytogenes comprising the lipopeptides, where the CTL epitope is from
XX an influenza virus or hepatitis C virus or Listeria monocytogenes protein
XX; (6) treating or preventing cancer; and (7) a vaccine against cancer
XX comprising the lipopeptide, where the CTL epitope is a tumour-specific
XX CTL epitope. The lipopeptide has cytostatic and virucide activities, and
XX can be used in vaccine. The lipopeptide is useful for preparing a
XX composition for treating or preventing cancer, or hepatitis C virus or
XX influenza virus. The present sequence is used in the exemplification of
XX the present invention.
XX
XX Sequence 17 AA;
XX
XX Query Match 100.0%; Score 89; DB 8; Length 17;
XX Best Local Similarity 100.0%; Pred. No. 9.4e-09;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPVH 17
   |||||
Db 1 TRSRKQTSRLKNIPVH 17

```



```

RESULT 4
AAR83304
ID AAR83304 standard; protein; 662 AA.
XX AC
XX AAR83304;
XX 25-MAR-2003 (revised)
XX DT 29-MAY-1996 (first entry)
XX DT
XX DE Canine distemper virus fusion glycoprotein.
XX KW
XX KW Avian influenza virus; rabies; canine distemper virus; CDV; measles;
XX KW antibody; vaccine; therapy; TROVAC.
XX OS Synthetic.
XX FN W09527780-A1.
XX PD 19-OCT-1995.
XX PF 06-APR-1995; 95WO-US004394.
XX PR 06-APR-1994; 94US-00224657.
XX PR 05-APR-1995; 95US-00416646.
XX PA (VIRO-) VIROGENETICS CORP.
XX PI Paoletti E, Tartaglia J, Taylor J, Gettig R;
XX WPI; 1995-366385/47.
XX DR N-PSDB; AAT00521.
XX DR
XX PT New modified recombinant viruses - contg. a canine distemper virus
XX PT antigen or measles virus M or N antigen in a non-essential region.
XX PS Example 15; Fig 15; 194pp; English.
XX CC This sequence represents the H6 promoted canine distemper virus (CDV)
XX CC fusion glycoprotein sequence. The DNA encoding this sequence is contained
XX CC in the plasmid pSDCDVHA. The encoding sequence was removed from this
XX CC plasmid and inserted into the C5 locus of another plasmid. The vector can
XX CC then be used to introduce the exogenous DNA into a fowlpox virus. This
XX CC procedure can be performed using exogenous DNA encoding a measles virus M
XX CC or N antigen. The recombinant viruses can then be used to induce an
XX CC antigenic or immunological response. The expression products of these
XX CC recombinants, and the antibodies produced can be used in binding assays
XX CC to determine the presence of CDV or measles virus in a sample. The
XX CC attenuated virulence of the viruses reduces the possibility of a runaway
XX CC infection due to the vaccination, in a vaccinated individual. The viruses
XX CC also reduce the transmission from vaccinated to unvaccinated individuals,
XX CC correct PR field.)
XX SQ Sequence 662 AA;
XX Query Match 100.0%; Score 89; DB 2; Length 662;
XX Best Local Similarity 100.0%; Pred. No. 6.5e-07;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 TRSRKQTSRLKNIPVH 17
XX DB |||||
XX 64 TRSRKQTSRLKNIPVH 80
XX
XX RESULT 5
XX AAB08102
XX ID AAB08102 standard; protein; 662 AA.
XX AC
XX AC AAB08102;
XX XX
XX DT 04-DEC-2000 (first entry)
XX DE Amino acid sequence of a fusion protein of canine distemper virus.
XX
XX Query Match 100.0%; Score 89; DB 2; Length 662;
XX Best Local Similarity 100.0%; Pred. No. 6.5e-07;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 TRSRKQTSRLKNIPVH 17
XX DB |||||
XX 64 TRSRKQTSRLKNIPVH 80
XX
XX RESULT 6
XX AAM50124
XX ID AAM50124 standard; protein; 662 AA.
XX AC
XX AC AAM50124;
XX XX
XX DT 21-DEC-2001 (first entry)
XX DE
XX DE Canine distemper virus recombinant fusion protein PCDVVF662.
XX KW
XX KW CDV; fusion protein; PCDVVF662; antigen; immune status;
XX KW vaccination status; dog.
XX OS
XX OS Canine distemper virus.
XX PN
XX PN W0200166568-A2.
XX PD
XX PD 13-SEP-2001.
XX PF
XX PF 07-MAR-2001; 2001WO-US007251.
XX PR
XX PR 09-MAR-2000; 2000US-00521738.
XX XX
XX XX (HESK-) HESKA CORP.
XX PA (COLS ) UNIV COLORADO STATE RES FOUND.

```

```

XX T helper cell epitope; CDV; immune response; canine vaccine.
XX OS
XX OS Canine distemper virus.
XX PN
XX PN W0200046390-A1.
XX PD
XX PD 10-AUG-2000.
XX PF
XX PF 07-FEB-2000; 2000WO-AU000070.
XX PR
XX PR 05-FEB-1999; 99AU-00008533.
XX PR 04-AUG-1999; 99AU-00002013.
XX XX
XX XX (UYME ) UNIV MELBOURNE.
XX PA (CSLC-) CSL LTD.
XX PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX PI Jackson DC, Souravi G, Walker J;
XX WPI; 2000-532904/48.
XX PT Novel T helper cell epitopes derived from canine distemper virus useful
XX PT for preparation of canine vaccines.
XX PS Example 1; Fig 1; 54pp; English.
XX CC The present sequence represents a fusion protein of canine distemper
XX CC virus (CDV). The protein was used to identify peptides AAB08076-B08101,
XX CC which are T helper cell epitopes. Compositions comprising these T cell
XX CC helper epitopes are useful for inducing an immune response in an animal.
XX CC The epitopes are useful as components of animal, in particular, canine
XX CC vaccines, either simply as synthetic peptide based vaccines and as
XX CC additions to vaccines containing more complex antigens
XX SQ Sequence 662 AA;
XX Query Match 100.0%; Score 89; DB 3; Length 662;
XX Best Local Similarity 100.0%; Pred. No. 6.5e-07;
XX Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 TRSRKQTSRLKNIPVH 17
XX DB |||||
XX 64 TRSRKQTSRLKNIPVH 80
XX
XX RESULT 6
XX AAM50124
XX ID AAM50124 standard; protein; 662 AA.
XX AC
XX AC AAM50124;
XX XX
XX DT 21-DEC-2001 (first entry)
XX DE
XX DE Canine distemper virus recombinant fusion protein PCDVVF662.
XX KW
XX KW CDV; fusion protein; PCDVVF662; antigen; immune status;
XX KW vaccination status; dog.
XX OS
XX OS Canine distemper virus.
XX PN
XX PN W0200166568-A2.
XX PD
XX PD 13-SEP-2001.
XX PF
XX PF 07-MAR-2001; 2001WO-US007251.
XX PR
XX PR 09-MAR-2000; 2000US-00521738.
XX XX
XX XX (HESK-) HESKA CORP.
XX PA (COLS ) UNIV COLORADO STATE RES FOUND.

```

XX Jensen WA, Lappin MR, Rosen DK, Andrews JS;  
PI WPI; 2001-639000/73.  
XX N-PSDB; AAH27071.  
XX Determining immune status or vaccination status of an animal to e.g.  
PT calicivirus comprises using a recombinant viral antigen.  
XX  
PS Claim 4; Page 129-131; 132pp; English.  
XX  
XX The present sequence is that of canine distemper virus fusion protein  
CC recombinant antigen, PCDFV662. The recombinant antigen, fused to an N-  
CC terminal His tag, was produced in Escherichia coli cells transformed by a  
CC recombinant vector comprising nucleic acid NCDVFL1986 (see AAH27071).  
CC PCDFV662 is an example of a recombinant infectious agent antigen status of  
CC be used in the method of the invention to determine the immune status of  
CC an animal. The method involves contacting a biological specimen of an  
CC animal (cat, dog or horse) with a recombinant antigen, and detecting the  
CC presence or absence of a complex between the recombinant antigen and an  
CC antibody present in the sample. The method determines whether the animal  
CC is protected against disease or should be vaccinated. Recombinant  
CC antigens (see AAM50107-24), nucleic acids encoding them (see AAH27054-  
CC 71), methods of producing them, and assay methods are provided  
XX  
SQ Sequence 662 AA;

Query Match 100.0%; Score 89; DB 4; Length 662;  
Best Local Similarity 100.0%; Pred. No. 6.5e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17  
DB 64 TRSRKQTSRLKNIPVH 80

RESULT 7  
ADM47653  
ID AAM47653 standard; protein; 662 AA.  
XX AC  
XX AAM47653;  
XX  
DT 21-FEB-2002 (first entry)  
XX  
DE Canine distemper virus, CDV, F.  
XX  
XX Virucide; vaccine; virus; virulence; canine distemper virus; CDV;  
XX measles; dog.  
XX  
OS Canine distemper virus.  
XX  
XX US6309647-B1.  
XX  
XX 30-OCT-2001.  
XX  
XX 15-JUL-1999; 99US-00354138.  
XX  
XX 15-JUL-1999; 99US-00354138.  
XX  
XX (AVET ) AVENTIS PASTEUR.  
XX  
XX Paolletti E, Tartaglia J, Taylor J, Gettig R;  
XX  
XX WPI; 2002-040232/05.  
XX  
XX N-PSDB; AB198922.  
XX  
XX Novel virus, useful for inducing immune response in dog against CDV,  
PT comprises the modified recombinant virus having attenuated virulence  
PT comprising exogenous DNA sequences encoding antigens of canine distemper  
XX virus (CDV) or measles virus.  
XX  
XX Example 15; Fig 15; 147pp; English.  
XX

CC The present invention relates to modified recombinant viruses, comprising  
CC inactivated virus-encoded genetic functions so that the viruses have  
CC attenuated virulence, yet retained efficiency. The viruses can contain  
CC DNA encoding a canine distemper virus (CDV) antigen or measles M or N  
CC antigen. The recombinant viruses are useful for inducing an antigenic or  
CC immunological response in a dog or other carnivore against CDV. The  
CC present sequence was used in an example from the present invention  
XX  
SQ Sequence 662 AA;

Query Match 100.0%; Score 89; DB 5; Length 662;  
Best Local Similarity 100.0%; Pred. No. 6.5e-07;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRSRKQTSRLKNIPVH 17  
DB 64 TRSRKQTSRLKNIPVH 80

RESULT 8  
ADM66120  
ID ADM66120 standard; protein; 662 AA.  
XX AC  
XX ADM66120;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Vaccinia virus H6-promoted CDV Fusion (F) protein #1.  
XX  
XX Canine distemper virus; NYVAC; TROVAC; ALVAC; rabies glycoprotein G;  
XX rabies; vaccine.  
XX  
OS Canine distemper virus; Onderstepoort strain.  
XX  
XX US2003082204-A1.  
XX  
XX 01-MAY-2003.  
XX  
XX 13-SEP-2001; 2001US-00951061.  
XX  
XX 20-NOV-1990; 90US-00621614.  
XX  
XX 07-MAR-1991; 91US-00666056.  
XX  
XX 11-JUN-1991; 91US-00713967.  
XX  
XX 22-OCT-1991; 91US-00776867.  
XX  
XX 06-MAR-1992; 92US-00847951.  
XX  
XX 31-AUG-1992; 92US-00938283.  
XX  
XX 08-JUN-1993; 93US-00073962.  
XX  
XX 12-AUG-1993; 93US-00105483.  
XX  
XX 06-APR-1994; 94US-00224657.  
XX  
XX 15-JUL-1999; 99US-00354138.  
XX  
XX (AVET ) AVENTIS PASTEUR.  
XX  
XX Paolletti E, Tartaglia J, Taylor J, Gettig R;  
XX  
XX WPI; 2003-567445/53.  
XX  
XX N-PSDB; ADM66154.  
XX  
XX New recombinant viruses comprising exogenous DNA encoding rabies  
PT glycoprotein G useful for eliciting protective immunity against rabies  
PT virus in a carnivore.  
XX  
XX Example 15; SEQ ID NO 86; 93pp; English.  
XX  
XX The invention relates to recombinant vaccinia and canarypox viruses  
CC comprising exogenous DNA encoding rabies glycoprotein G in a nonessential  
CC region of the virus genome. Also included are a recombinant vaccinia virus  
CC (comprising exogenous DNA encoding rabies glycoprotein G in a  
CC nonessential region of the virus genome, where at least one open reading  
CC frame (ORF) selected from J2, B13, +B14R, A26L, 156R, C7L-K1L, and 14L is  
CC deleted from the virus), a recombinant canarypox virus (produced by  
CC attenuation through multiple serial passages on chick embryo fibroblasts,  
CC subjecting a master seed from to successive plaque purifications under

agar and amplifying a plaque clone through multiple additional passages, where the virus contains exogenous DNA encoding rabies glycoprotein G in a nonessential region of the virus genome), inducing an antigenic or immunological response in a carnivore against rabies virus (by administering to the dog, cat or other carnivore a composition comprising the virus above in a mixture with a carrier), expressing a gene product in a cell cultured in vitro by introducing into the cell a virus of the invention. One or more (optionally all) ORFs selected from a thymidine kinase gene, a haemorrhagic region, an A type inclusion body, a haemagglutinin gene, a host range region, and a ribonucleotide reductase large subunit gene, may also be deleted. The attenuated Vaccinia virus is termed a NYVAC virus. The attenuated canarypox virus is termed ALVAC recombinant virus. Also included are attenuated fowl pox viruses termed TROVAC. The recombinant viruses are useful as vaccines for protecting a dog, cat or other carnivore against rabies. The modified recombinant viruses are effective as vaccines and are safer than some other recombinant viruses due to the deletion of genes affecting virulence that are not essential for virus growth in tissue culture. The present sequence is a Canine distemper virus protein expressed from an attenuated virus of the invention.

Query Match 662 AA;

Best Local Similarity 100.0%; Score 89; DB 7; Length 662;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRSRKQTSRLKNIPIVH 17  
| | | | | | | | | | | | | | | | | |  
Db 64 TRSRKQTSRLKNIPIVH 80

RESULT 9  
ABG00101  
ID ABG00101 standard; protein; 280 AA.

AC ABG00101;  
XX  
XX 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #92.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS64288.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 20; SEQ ID NO 30460; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used

CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 280 AA;

Query Match 50.0%; Score 44.5; DB 4; Length 280;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 2 RSRKQTSRHR-LKNIPV 16  
| | | | | | | | | | | | | | | | | |  
Db 230 RSRKQTSRHR-LKNIPV 245

RESULT 10  
ABG00285

ID ABG00285 standard; protein; 280 AA.

AC ABG00285;

XX 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #276.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS64472.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX Claim 20; SEQ ID NO 30644; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 280 AA;

Query Match 50.0%; Score 44.5; DB 4; Length 280;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 2 RSRKQTSR-LKNIPV 16  
DB 230 RSRQTNHRQMGKIPV 245

RESULT 11  
ABP40732  
ID ABP40732 standard; protein; 103 AA.  
XX AC ABP40732;  
XX DT 24-JUL-2002 (first entry)  
XX DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5577.  
XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
XX KW antibacterial; gene therapy.  
XX OS Staphylococcus epidermidis.  
XX PN US6380370-B1.  
XX PD 30-APR-2002.  
XX PF 13-AUG-1998; 98US-00134001.  
XX PR 14-AUG-1997; 97US-0055779P.  
XX PR 08-NOV-1997; 97US-0064964P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Doucette-Stamm LA, Bush D;  
XX WPI; 2002-381255/41.  
XX DR N-PSDB; ABN93277.  
XX Novel isolated nucleic acid encoding a Staphylococcus epidermis polypeptide, useful for diagnosing and treating bacterial infections.  
XX Disclosure; SEQ ID NO 5577; 267pp; English.  
XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP35124 to ABP37960. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection. N.B. The sequence data for this patent did not form part of the printed specification, but was obtained

CC in electronic format directly from the USPTO web site  
XX Sequence 103 AA;  
SQ Query Match 49.4%; Score 44; DB 5; Length 103;  
Best Local Similarity 69.2%; Pred. No. 9.1;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 KQTSRLKNIPVH 17  
DB 81 KQSDRLKEIPVH 93

RESULT 12  
ADS08038  
ID ADS08038 standard; protein; 103 AA.  
XX AC ADS08038;  
XX DT 04-NOV-2004 (first entry)  
XX DE Staphylococcus epidermis polypeptide seqid 7333.  
XX KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;  
XX KW recombinant expression vector; infection; computer readable medium;  
XX KW computer based system.  
XX OS Staphylococcus epidermidis.  
XX PN US2004147734-A1.  
XX PD 29-JUL-2004.  
XX PF 01-DEC-2003; 2003US-00724972.  
XX PR 08-NOV-1997; 97US-0064964P.  
XX PR 13-AUG-1998; 98US-00134001.  
XX PR 29-NOV-1999; 99US-00450969.  
XX PA (DOUC/) DOUCETTE-STAMM L.  
XX PA (BUSH/) BUSH D.  
XX PI Doucette-Stamm L, Bush D;  
XX WPI; 2004-580138/56.  
XX DR N-PSDB; ADS04266.  
XX New isolated polypeptide and encoding nucleic acid derived from Staphylococcus epidermidis, useful for diagnosing, preventing and/or treating an S. epidermidis bacterial infection.  
XX Claim 17; SEQ ID NO 7333; 741pp; English.

The invention describes an isolated nucleic acid comprising a nucleotide sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO: 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as given in the specification. Also described are: a recombinant expression vector; a cell comprising a recombinant expression vector of (1); producing an S. epidermidis polypeptide; an isolated nucleic acid comprising a nucleotide sequence of at least 8 nucleotides in length; a vaccine composition for prevention or treatment of an S. epidermidis infection, comprising a nucleic acid cited above and a carrier; treating a subject for S. epidermidis infection; a recombinant or substantially pure preparation of an S. epidermidis polypeptide or its fragment; a vaccine composition for prevention or treatment of an S. epidermidis infection; detecting the presence of a Staphylococcus nucleic acid in a sample; a computer readable medium having recorded in it the nucleotide sequences with SEQ ID NO: 1-3772 or its fragments; a computer based system for identifying fragments of the Staphylococcus genome of commercial importance; a computer based system for identifying fragments of the Staphylococcus plasmids of commercial importance; identifying commercially important nucleic acid fragments of the Staphylococcus



Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RSRKQTSRLKKNIP 15  
 | : | | | | | : |  
 Db 140 RNTKQTIHLKRLP 153

RESULT 15  
 ADS21233  
 ID ADS21233 standard; protein; 354 AA.  
 XX AC  
 AC ADS21233;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Bacterial polypeptide #10266.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX  
 OS Bacteria.

CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at seqdata.uspto.gov/sequence.html.  
 XX  
 SQ Sequence 354 AA;  
 Query Match 47.2%; Score 42; DB 8; Length 354;  
 Best Local Similarity 46.7%; Pred. No. 87;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 SRKQTSRLKKNIPVH 17  
 | : | | | | : | |  
 Db 324 SKKETQIRIANPPAH 338

Search completed: September 28, 2005, 17:50:23  
 Job time : 93.25 secs

US2003233675-A1.  
 18-DEC-2003.  
 20-FEB-2003; 2003US-00369493.  
 21-FEB-2002; 2002US-0360039P.  
 (CAOV/) CAO Y.  
 (HINK/) HINKLE G J.  
 (SLAT/) SLATER S C.  
 (CHEN/) CHEN X.  
 (GOLD/) GOLDMAN B S.  
 Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
 WPI; 2004-061375/06.  
 New recombinant DNA construct comprising a promoter positioned to provide  
 for expression of a polynucleotide encoding a polypeptide from a  
 microbial source, useful for producing plants with improved properties.  
 Claim 1; SEQ ID NO 10266; 122pp; English.  
 The invention relates to a recombinant DNA construct comprising a  
 promoter functional in a plant cell, where the promoter is positioned to  
 provide for expression of a polynucleotide encoding a polypeptide from a  
 microbial source. The invention also relates to a transformed plant  
 comprising the recombinant DNA construct and a method of producing a  
 transformed plant having an improved property. The plant is a crop plant  
 such as maize or soybean. The method of producing a transformed plant  
 having an improved property comprises transforming a plant with the  
 recombinant DNA construct and growing the transformed plant, where the  
 polynucleotide or polypeptide is useful for improving plant properties.  
 The recombinant DNA construct is useful for producing plants with  
 improved plant properties, e.g. improved cold, heat or drought tolerance,  
 tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 increased resistance to plant disease, better growth rate by modification  
 of the cell cycle pathway with plant growth regulators, increased rate of  
 homologous recombination, modified seed oil or protein yield and/or  
 content, improved yield by modification of carbohydrate, nitrogen or  
 phosphorus use and/or uptake, by modification of photosynthesis or by  
 providing improved plant growth and development under at least one stress  
 condition, improved lignin production or improved galactomannan  
 production. This sequence represents a bacterial polypeptide used in the  
 scope of the invention. Note: The sequence data for this patent did not

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:37:01 ; Search time 91.25 Seconds  
(without alignments)  
72.054 Million cell updates/sec

Title: US-10-705-819B-4

Perfect score: 89

Sequence: 1 PRTSDRPVSYTMNTRTS 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	17	AA08079	Amino aci
2	89	100.0	17	ADK00505	Immunogen
3	89	100.0	17	ADJ84579	T-helper
4	89	100.0	662	AA083304	Canine di
5	89	100.0	662	AA080102	Amino aci
6	89	100.0	662	AA050124	Canine di
7	89	100.0	662	AA047653	Canine di
8	89	100.0	662	AD066120	Vaccinia
9	45	50.6	331	AB059547	Drosophila
10	45	50.6	402	AB025408	Aspergill
11	43	48.3	155	AB081533	Pseudomon
12	42.5	47.8	181	AA001841	Human gen
13	42.5	47.8	181	AB064163	Human alb
14	42.5	47.8	181	ADL77428	Albumin f
15	42.5	47.8	486	AD065683	Human pro
16	42.5	47.8	486	ADJ70013	Human hea
17	42.5	47.8	655	AD065930	Novel hum
18	42.5	47.8	778	AA001791	Human gen
19	42.5	47.8	778	AB064165	Human alb
20	42.5	47.8	778	ADL77430	Albumin f
21	42.5	47.8	784	AD065924	Novel hum
22	42	47.2	529	ABJ26367	Aspergill
23	42	47.2	581	AB068988	Pseudomon
24	42	47.2	614	AA017907	Pseudomon
25	42	47.2	614	AA017906	Pseudomon

26	42	47.2	614	2	AA017905	Pseudomon
27	42	47.2	614	2	AA017908	Pseudomon
28	42	47.2	614	2	AA017909	Pseudomon
29	42	47.2	614	2	AA017904	Pseudomon
30	42	47.2	626	6	ABU03104	Alpha amy
31	42	47.2	632	2	AA015470	Maltopent
32	42	47.2	820	6	AB053012	Human spl
33	42	47.2	1310	7	ADK40915	Novel hum
34	42	47.2	1310	8	AD015638	Kinase 41
35	42	47.2	1373	7	ADC31009	Human nov
36	42	47.2	1520	8	AD065659	Novel hum
37	42	47.2	1999	6	ABR43628	Mouse CLA
38	42	47.2	2090	4	AB099495	Human CLA
39	42	47.2	2090	5	ABG94416	Human CLA
40	42	47.2	2090	5	ABG61705	Human cad
41	42	47.2	2090	6	ABR43635	Human CLA
42	41	46.1	78	4	AAU61202	Propionib
43	41	46.1	78	6	ABM57721	Propionib
44	41	46.1	114	3	ABG58274	Arabidops
45	41	46.1	133	7	ADJ69112	Human hea

## ALIGNMENTS

## RESULT 1

AA08079  
ID AA08079 standard; peptide; 17 AA.

XX AA08079;

AC AA08079;

DT 04-DEC-2000 (first entry)

DE Amino acid sequence of a helper T cell epitope from CDV.

XX T helper cell epitope; CDV; immune response; canine vaccine.

XX Canine distemper virus.

OS WO200046390-A1.

XX 10-AUG-2000.

XX 07-FEB-2000; 2000WO-AU000070.

XX 05-FEB-1999; 99AU-00008533.

XX 04-AUG-1999; 99AU-00002013.

XX (UYME ) UNIV MELBOURNE.

XX (CSLC-) CSL LTD.

XX (CSIR ) COMMONWEALTH SCI & IND RES ORG.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX Jackson DC, Souravi G, Walker J;

XX WPI; 2000-532904/48.

XX Novel T helper cell epitopes derived from canine distemper virus useful for preparation of canine vaccines.

XX Claim 1; Page 28; 54pp; English.

XX AA08076-B08101 represent T helper cell epitopes, derived from canine distemper virus (CDV). Compositions comprising these T cell helper epitopes are useful for inducing an immune response in an animal. The epitopes are useful as components of animal, in particular, canine vaccines, either simply as synthetic peptide based vaccines and as additions to vaccines containing more complex antigens

XX SQ Sequence 17 AA;

XX Query Match 100.0%; Score 89; DB 3; Length 17;





## RESULT 4

AAR83304  
ID AAR83304 standard; protein; 662 AA.

AC AAR83304;

	25-MAR-2003	(revised)
DT	29-MAY-1996	(first entry)

DE Canine distemper virus fusion glycoprotein.

KW Avian influenza virus; rabies; canine distemper virus; CDV; measles;  
 KW antibody; vaccine; therapy; TROVAC.

OS Synthetic.

PN W09527780-A1.

19-OCT-1995.

XX  
PF 06-APR-1995; 95WO-US004394.AA  
PR 06-APR-1994; 94US-00224657.

PR 05-APR-1995; 95US-00416646.

PA (VIRO-) VIROGENETICS CORP.

PI Paoletti E, Tartaglia J, Taylor J, Gettig R;

WPI: 1995-366385/47.

DR N-PSDB; AAT00521;

PT New modified recombinant viruses - contg. a canine distemper virus antigen or measles virus M or N antigen in a non-essential region.

PS Example 15: Fig 15: 194pp: English.

This sequence represents the H6 promoted canine distemper virus (CDV) fusion glycoprotein sequence. The DNA encoding this sequence is contained in the plasmid pSPCDVHA. The encoding sequence was removed from this plasmid and inserted into the C5 locus of another plasmid. The vector can then be used to introduce the exogenous DNA into a fowlpox virus. This procedure can be performed using exogenous DNA encoding a measles virus M or N antigen. The recombinant viruses can then be used to induce an antigenic or immunological response. The expression products of these recombinants, and the antibodies produced can be used in binding assays to determine the presence of CDV or measles virus in a sample. The attenuated virulence of the viruses reduces the possibility of a runaway infection due to the vaccination, in a vaccinated individual. The viruses also reduce the transmission from vaccinated to unvaccinated individuals, and reduces environmental contamination. (Updated on 25-MAR-2003 to correct PR field.)

Sequence 662 AA:

Query Match	100.0%;	Score 89;	DB 2;	Length 662;
Best Local Similarity	100.0%;	Pred. No. 1.1e-05;		
Matches 17;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

1 PRTSDRPPVSYTMNRTS 17

db 50 PRTSDRPVSYTMNRTS 66

## RESULT 5

AAB08102  
ID AAB08102 standard; protein; 662 AA.

AC AAB08102;

04-DEC-2000 (first entry)

DE Amino acid sequence of a fusion protein of canine distemper virus.

```

XX Jensen WA, Lappin MR, Rosen DK, Andrews JS;
XX WPI; 2001-639000/73.
XX N-PSDB; AAH27071.
XX Determining immune status or vaccination status of an animal to e.g.
XX PT calicivirus comprises using a recombinant viral antigen.
XX PS Claim 4; Page 129-131; 132pp; English.
XX The present sequence is that of canine distemper virus fusion protein
XX recombinant antigen, PCDVf662. The recombinant antigen, fused to an N-
XX terminal His tag, was produced in Escherichia coli cells transformed by a
XX recombinant vector comprising nucleic acid NCDVFI986 (see AAH27071).
XX PCDVf662 is an example of a recombinant infectious agent antigen status of
XX an animal. The method involves contacting a biological specimen of an
XX animal (cat, dog or horse) with a recombinant antigen, and detecting the
XX presence or absence of a complex between the recombinant antigen and an
XX antibody present in the sample. The method determines whether the animal
XX is protected against disease or should be vaccinated. Recombinant
XX antigens (see AAM50107-24), nucleic acids encoding them (see AAH27054-
XX 71), methods of producing them, and assay methods are provided
XX SQ Sequence 662 AA;

Query Match 100.0%; Score 89; DB 4; Length 662;
Best Local Similarity 100.0%; Pred. NO. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
DB 50 PRSDRPVSYTMNRTS 66

RESULT 7
AAM47653
ID AAM47653 standard; protein; 662 AA.
AC AAM47653;
XX 21-FEB-2002 (first entry)
XX Canine distemper virus, CDV, F.
XX Virucide; vaccine; virus; virulence; canine distemper virus; CDV;
XX measles; dog.
XX Canine distemper virus.
XX US6309647-B1.
XX 30-OCT-2001.
XX 15-JUL-1999; 99US-00354138.
XX 15-JUL-1999; 99US-00354138.
XX (AVET ) AVENTIS PASTEUR.
XX Paolletti E, Tartaglia J, Taylor J, Gettig R;
XX WPI; 2002-040232/05.
XX N-PSDB; AB198922.
XX Novel virus, useful for inducing immune response in dog against CDV,
XX PT comprises the modified recombinant virus having attenuated virulence
XX PT comprising exogenous DNA sequences encoding antigens of canine distemper
XX PT virus (CDV) or measles virus.
XX Example 15; Fig 15; 147pp; English.
XX

The present invention relates to modified recombinant viruses, comprising
inactivated virus-encoded genetic functions so that the viruses have
attenuated virulence, yet retained efficiency. The viruses can contain
DNA encoding a canine distemper virus (CDV) antigen or measles M or N
antigen. The recombinant viruses are useful for inducing an antigenic or
immunological response in a dog or other carnivore against CDV. The
present sequence was used in an example from the present invention
SQ Sequence 662 AA;

Query Match 100.0%; Score 89; DB 5; Length 662;
Best Local Similarity 100.0%; Pred. NO. 1.1e-05;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRSDRPVSYTMNRTS 17
DB 50 PRSDRPVSYTMNRTS 66

RESULT 8
ADM66120
ID ADM66120 standard; protein; 662 AA.
XX AC ADM66120;
XX 03-JUN-2004 (first entry)
XX Vaccinia virus H6-promoted CDV Fusion (F) protein #1.
XX Canine distemper virus; NYVAC; TROVAC; ALVAC; rabies glycoprotein G;
XX rabies; vaccine.
XX Canine distemper virus; Onderstepoort strain.
XX US2003082204-A1.
XX 01-MAY-2003.
XX 13-SEP-2001; 2001US-00951061.
XX 20-NOV-1990; 90US-00621614.
XX 07-MAR-1991; 91US-00666056.
XX 11-JUN-1991; 91US-00713967.
XX 22-OCT-1991; 91US-00776867.
XX 06-MAR-1992; 92US-00847951.
XX 31-AUG-1992; 92US-00938283.
XX 08-JUN-1993; 93US-00073962.
XX 12-AUG-1993; 93US-00105483.
XX 06-APR-1994; 94US-00224657.
XX 15-JUL-1999; 99US-00354138.
XX (AVET ) AVENTIS PASTEUR.
XX Paolletti E, Tartaglia J, Taylor J, Gettig R;
XX WPI; 2003-567445/53.
XX N-PSDB; ADM66154.
XX New recombinant viruses comprising exogenous DNA encoding rabies
XX glycoprotein G useful for eliciting protective immunity against rabies
XX virus in a carnivore.
XX Example 15; SEQ ID NO 86; 93pp; English.
XX

The invention relates to recombinant vaccinia and canarypox viruses
comprising exogenous DNA encoding rabies glycoprotein G in a nonessential
region of the virus genome. Also included are a recombinant vaccinia virus
comprising exogenous DNA encoding rabies glycoprotein G in a
nonessential region of the virus genome, where at least one open reading
frame (ORF) selected from J2, B13, +B14R, A26L, 156R, C7L-K1L, and 14L is
deleted from the virus), a recombinant canarypox virus (produced by
attenuation through multiple serial passages on chick embryo fibroblasts,
subjecting a master seed from to successive plaque purifications under

```

CC agar and amplifying a plaque clone through multiple additional passages,  
 CC where the virus contains exogenous DNA encoding rabies glycoprotein G in  
 CC a nonessential region of the virus genome), inducing an antigenic or  
 CC immunological response in a carnivore against rabies virus (by  
 CC administering to the dog, cat or other carnivore a composition comprising  
 CC the virus above in a mixture with a carrier), expressing a gene product  
 CC in a cell cultured in vitro by introducing into the cell a virus of the  
 CC invention. One or more (optionally all) ORFs selected from a thymidine  
 CC kinase gene, a haemorrhagic region, an A type inclusion body, a  
 CC haemagglutinin gene, a host range region, and a ribonucleotide reductase  
 CC large subunit gene, may also be deleted. The attenuated Vaccinia virus is  
 CC termed a NYVAC virus. The attenuated canarypox virus is termed ALVAC  
 CC recombinant virus. Also included are attenuated fowl pox viruses termed  
 CC TROVAC. The recombinant viruses are useful as vaccines for protecting a  
 CC dog, cat or other carnivore against rabies. The modified recombinant  
 CC viruses are effective as vaccines and are safer than some other  
 CC recombinant viruses due to the deletion of genes affecting virulence that  
 CC are not essential for virus growth in tissue culture. The present  
 CC sequence is a Canine distemper virus protein expressed from an attenuated  
 CC virus of the invention.

XX  
 SQ Sequence 662 AA;

Query Match 100.0%; Score 89; DB 7; Length 662;  
 Best Local Similarity 100.0%; Pred. NO. 1.1e-05; Mismatches 0; Indels 0; Gaps 0;  
 Matches 17; Conservative 0;

Qy 1 PRTSDRPVSYTMNRTR 17  
 |||||  
 Db 50 PRTSDRPVSYTMNRTR 66

RESULT 9  
 ABB59547  
 ID ABB59547 standard; protein; 331 AA.

XX  
 AC ABB59547;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 5433.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL03650.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 5433; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU16176-ABU130511), expressed DNA  
 CC sequences (ABU16175) and the encoded proteins (ABU57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 331 AA;

Query Match 50.6%; Score 45; DB 4; Length 331;  
 Best Local Similarity 56.2%; Pred. No. 73;  
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTR 16  
 | : |||||  
 Db 204 PELIQRELSYTMNRIR 219

RESULT 10  
 ABB25408  
 ID ABB25408 standard; protein; 402 AA.

XX  
 AC ABB25408;

XX 16-APR-2003 (first entry)

XX Aspergillus fumigatus essential gene protein #66.

XX Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;  
 KW cancer; contamination; biofilm; antibody; immune response.

XX Aspergillus fumigatus.

XX WO200286090-A2.

XX 31-OCT-2002.

XX 23-APR-2002; 2002WO-US013142.

XX 23-APR-2001; 2001US-0285697P.

XX 27-APR-2001; 2001US-0287066P.

XX 05-JUN-2001; 2001US-0295890P.

XX 09-JUL-2001; 2001US-0303899P.

XX 31-AUG-2001; 2001US-0316362P.

XX (ELIT-) ELITRA PHARM INC.

XX Jiang B, Tishkoff D, Zamudio C, Eroehkin AM, Hu W, Lemieux SM;

XX WPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of *Aspergillus*  
 PT fumigatus, useful for treating or preventing infections by *A. fumigatus*,  
 PT or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure; Page; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as *A. fumigatus*, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or contain contamination of an object  
 CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for  
 CC expressing recombinant protein for characterisation, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*  
 CC *fumigatus* to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an



XX Claim 11; Page 532-533; 567pp; English.

PS AAD05579-AA05658 represent cDNAs corresponding to 28 human secreted

CC protein genes and AAE01770-AAE01849 represent the proteins they encode.

CC AAE01850-AAE01860 represent human secreted protein fragments or variants.

CC The genes and their secreted proteins are useful for preventing, treating

CC or ameliorating medical conditions, e.g., by protein or gene therapy.

CC Pathological conditions can be diagnosed by determining the amount of the

CC new protein in a sample or by determining the presence of mutations in

CC the new genes. Specific uses are described for each of the 28 genes,

CC based on the tissues in which they are most highly expressed, and include

CC developing products for the diagnosis or treatment of proliferative

CC disorders, cancer, tumours, foetal and developmental abnormalities,

CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune

CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,

CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),

CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,

CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,

CC angiogenic disorders, kidney disorders, gastrointestinal disorders,

CC pregnancy-related disorders, endocrine disorders, and infections. The

CC proteins can also be used to aid wound healing and epithelial cell

CC proliferation, to prevent skin aging due to sunburn, to maintain organs

CC before transplantation, for supporting cell culture of primary tissues,

CC to regenerate tissues, to identify their cognate ligands or binding

CC partners, and in chemotaxis, and can be used as a food additive or

CC preservative to modify storage properties. Antibodies specific for a

CC protein of the invention can be used in alleviating symptoms associated

CC with the disorders mentioned above, and in diagnostic immunoassays e.g.

CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The

CC present sequence represents a human secreted protein of the invention

XX Sequence 181 AA;

Qy Query Match 47.8%; Score 42.5; DB 4; Length 181;

Db Best Local Similarity 56.2%; Pred. No. 99;

Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 PRTSDRP-VSYTMNRT 15

Db 98 PRTSDKPHIRPVLNRT 113

RESULT 13

ID ABG64163 standard; protein; 181 AA.

AC ABG64163;

XX 27-AUG-2002 (first entry)

DE Human albumin fusion protein #838.

XX Albumin fusion protein, therapeutic protein X; human albumin; HA;

KW human serum albumin; HSA; cancer; reproductive disorder;

KW digestive disorder; immune disorder; endocrine disorder;

KW haematopoietic disorder; neural disorder; connective disorder;

KW cytostatic; antiinfertility; antiinflammatory; antiulcer;

KW immunomodulator; anti-HIV; antidiabetic; haemostatic; neurotropic;

KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;

KW osteopathic; antiarthritic.

XX Homo sapiens.

OS Synthetic.

XX WO200177137-A1.

PN 18-OCT-2001.

XX 12-APR-2001; 2001WO-US011988.

PF 12-APR-2000; 2000US-0229358P.

XX 25-APR-2000; 2000US-0199384P.

PR

PR 21-DEC-2000; 2000US-0256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI; 2002-010886/01.

XX New fusion protein for treating disease e.g. diabetes comprises an

PT albumin fused to a therapeutic protein.

XX Claim 1; Page 1058-1059; 2102pp; English.

XX The present invention relates to albumin fusion proteins comprising a

CC therapeutic protein X and human albumin (HA, also known as human serum

CC albumin, HSA). The proteins are useful for treating a disease or disorder

CC that may be modulated by therapeutic protein X. The albumin extends the

CC shelf-life of protein X, and may increase its biological in vitro/in vivo

CC activity. The protein is useful for treating and diagnosing disorders

CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's

CC disease, ulcerative colitis), immune disorders (e.g. acquired

CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),

CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,

CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,

CC schizophrenia) and connective disorders (e.g. osteoporosis, arthritis).

CC ABG63326-ABG65518 represent albumin fusion proteins of the invention

XX Sequence 181 AA;

Qy Query Match 47.8%; Score 42.5; DB 5; Length 181;

Db Best Local Similarity 56.2%; Pred. No. 99;

Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 1 PRTSDRP-VSYTMNRT 15

Db 98 PRTSDKPHIRPVLNRT 113

RESULT 14

ID ADL77428 standard; protein; 181 AA.

AC ADL77428;

XX 20-MAY-2004 (first entry)

DE Albumin fusion protein related therapeutic protein X, SRQ ID No 910.

XX albumin fusion protein; cytostatic; antianaemic; antiarthritic;

KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;

KW antipsoriatic; antibacterial; osteopathic; dermatological; antiulcer;

KW immunomodulator; antiarrhythmic; cardiac; neurotropic; antiparkinsonian; antipapillary;

KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;

KW antidiabetic; anabolic; hypertensive; vulnary; gene therapy; cancer;

KW reproductive system disorder; therapeutic protein.

XX Unidentified.

OS US2004010134-A1.

PN 15-JAN-2004.

XX 12-APR-2001; 2001US-00833245.

PF 12-APR-2000; 2000US-0229358P.

XX 25-APR-2000; 2000US-0199384P.

PR 21-DEC-2000; 2000US-0256931P.

XX (ROSE/) ROSEN C A.

PA (ROSE/) HASELTINE W A.

XX Rosen CA, Haseltine WA;

XX

DR WPI; 2004-090519/09.  
 XX New albumin fusion proteins, useful for diagnosing, treating, preventing  
 PT or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,  
 PT asthma, inflammatory bowel disease or Alzheimer's disease.  
 XX  
 PS Disclosure; SEQ ID NO 910; 279pp; English.  
 XX  
 CC The invention relates to a novel albumin fusion protein. The invention  
 CC further relates to: a composition comprising the albumin fusion protein  
 CC and a pharmaceutical carrier; a kit comprising the composition of the  
 CC albumin fusion protein formula; a method of treating a disease or  
 CC disorder in a patient comprising the step of administering the albumin  
 CC fusion protein; a method of treating a patient with a disease or disorder  
 CC that is modulated by Therapeutic protein: X, or its fragment or variant;  
 CC a method of extending the shelf life of Therapeutic protein: X, or its  
 CC fragment or variant; a nucleic acid molecule comprising a polynucleotide  
 CC sequence encoding the albumin fusion protein; a vector comprising the  
 CC nucleic acid molecule of the albumin fusion protein; and a host cell  
 CC comprising the nucleic acid molecule of the albumin fusion protein. The  
 CC albumin fusion protein and its compositions have the following  
 CC activities: cytostatic, antitumorigenic, antiarthritic, antidiabetic, anti-  
 CC HIV, immunosuppressive, antiinflammatory, antipruritic, antibacterial,  
 CC osteoprotective, dermatological, antitumor, immunomodulator, antiarrhythmic,  
 CC cardiact, neuroprotective, antiparkinsonian, tranquilizer, uropathic,  
 CC neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic,  
 CC hypertensive, and vulnerary. The albumin fusion protein nucleic acid may  
 CC be used in gene therapy to treat disorders. The albumin fusion protein is  
 CC useful for diagnosing, treating, preventing or ameliorating diseases or  
 CC disorders comprising indication: Y. The diseases or disorders include:  
 CC cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer),  
 CC immune or hematopoietic diseases (e.g. anemia, Hodgkin's disease, acute  
 CC lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,  
 CC autoimmune disease, inflammatory bowel disease, psoriasis or Lyme  
 CC disease), reproductive system disorders (e.g. prostatitis, inguinal  
 CC hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-  
 CC Leydig tumors), musculoskeletal diseases (e.g. giant cell tumors,  
 CC Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy  
 CC or cachexia), cardiovascular disease (e.g. rhabdomyoma, heart disease,  
 CC arrhythmia, cardiac arrest, heart valve disease, hypernatremia or  
 CC hyponatremia), mixed foetal diseases (e.g. foetal alcohol syndrome,  
 CC Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay  
 CC Sachs disease), excretory diseases (e.g. urinary incontinence, urinary  
 CC tract infections or renal disorders), neural or sensory disease (e.g.  
 CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,  
 CC cerebellar ataxia, attention deficit disorder, autism or obsessive  
 CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or  
 CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's  
 CC disease or glomerulonephritis), digestive diseases (e.g. portal  
 CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)  
 CC or connective tissue or epithelial diseases (e.g. Crohn's disease,  
 CC scleroderma, wound healing or epidermolysis bullosa). This sequence  
 CC represents a therapeutic protein X relating to the albumin fusion protein  
 CC of the invention. The sequence listing data for this specification was  
 CC downloaded from the USPTO website.  
 XX  
 SQ Sequence 181 AA;

Query Match 47.8%; Score 42.5; DB 8; Length 181;  
 Best Local Similarity 56.2%; Pred. No. 99;  
 Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 PRSDRP-VSYTWNRT 15  
 DB 98 PRSDRPHRPNVLRNT 113

RESULT 15  
 ADB65683  
 ID ADB65683 standard; protein; 486 AA.  
 XX  
 AC ADB65683;  
 XX

DT 04-DEC-2003 (first entry)  
 XX  
 DE Human protein encoded by clone TRACH20058000.  
 XX  
 KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;  
 KW cell regeneration; membrane protein; signal transduction-related protein;  
 KW transcription-related protein; osteoporosis; neurological disease;  
 KW cancer; tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1308459-A2.  
 XX  
 PD 07-MAY-2003.  
 XX  
 XX 28-MAR-2002; 2002EP-00007401.  
 PF  
 XX 05-NOV-2001; 2001JP-00379298.  
 PR  
 PR 25-JAN-2002; 2002US-00350978.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR WPI; 2003-450961/43.  
 DR N-PSDB; ADB63713.  
 XX  
 PT New polynucleotides and polypeptides, useful for developing a diagnostic  
 PT marker or medicines for regulation of their expression and activity, or  
 PT as targets of gene therapy.  
 XX  
 XX Claim 1; Page; 222pp; English.  
 PS  
 XX The invention discloses a polynucleotide comprising a sequence selected  
 CC from 1970 fully defined nucleotide sequences which encode novel  
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
 CC or its partial peptide, an antibody binding to the polypeptide or peptide  
 CC of the polynucleotide, immunologically assaying the polypeptide or  
 CC peptide of the polynucleotide by contacting the polypeptide or peptide  
 CC with the antibody of the encoded protein, and observing the binding  
 CC between the two, a transformant carrying the polynucleotide in an  
 CC expressible manner and an antisense polynucleotide. The oligonucleotide  
 CC is useful as a primer for synthesising the polynucleotide, or as a probe  
 CC for detecting the polynucleotide. The polynucleotides and encoded  
 CC proteins are useful as pharmaceutical agents and many disease-related  
 CC genes may be included in them, for developing a diagnostic marker or  
 CC medicines for regulation of their expression and activity, or as targets  
 CC of gene therapy. The genes are involved in tissue and/or cell  
 CC regeneration. Membrane proteins, signal transduction-related proteins,  
 CC transcription-related proteins, disease-related proteins and genes  
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate  
 CC the activity or expression of the encoded protein to treat diseases. The  
 CC sequence presented is a protein of the invention. Note: Some of the  
 CC sequence data for this patent is not represented in the printed  
 CC specification, but is based on sequence information supplied by the  
 CC European Patent Office.  
 XX  
 SQ Sequence 486 AA;

Query Match 47.8%; Score 42.5; DB 7; Length 486;  
 Best Local Similarity 56.2%; Pred. No. 2.7e+02;  
 Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 PRSDRP-VSYTWNRT 15  
 DB 52 PRSDRPHRPNVLRNT 67

Search completed: September 28, 2005, 17:50:21

Job time : 94.25 secs

---

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:56:03 ; Search time 87 Seconds

(without alignments)  
79.534 Million cell updates/sec

Title: US-10-705-819B-4

Perfect score: 89

Sequence: 1 PRSDRPVSYTMNRTS 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1826554 seqs, 407025358 residues

Total number of hits satisfying chosen parameters: 1826554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	17	17	US-10-935-155A-4
2	89	100.0	17	18	US-10-705-819B-4
3	89	100.0	662	10	US-09-951-061A-141
4	89	100.0	662	15	US-10-670-695-36
5	89	100.0	662	17	US-10-935-155A-27
6	89	100.0	662	18	US-10-705-819B-27
7	47	52.8	117	16	US-10-437-963-104337
8	47	52.8	302	16	US-10-437-963-134736
9	46	51.7	297	14	US-10-156-761-11575
10	45	50.6	94	16	US-10-437-963-168921
11	45	50.6	402	14	US-10-128-714-3066
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 141, App
					Sequence 36, Appl
					Sequence 27, Appl
					Sequence 27, Appl
					Sequence 104337,
					Sequence 134736,
					Sequence 11575, A
					Sequence 168921,
					Sequence 3066, Ap

12	44	49.4	49	16	US-10-425-115-341136
13	44	49.4	50	16	US-10-437-963-190524
14	44	49.4	79	16	US-10-437-963-126154
15	44	49.4	102	16	US-10-437-963-108772
16	44	49.4	113	16	US-10-437-963-185445
17	44	49.4	124	16	US-10-425-115-323905
18	43	48.3	111	16	US-10-437-963-151049
19	42.5	47.8	62	15	US-10-424-599-234131
20	42.5	47.8	181	11	US-09-833-245-910
21	42.5	47.8	486	15	US-10-104-047-3837
22	42.5	47.8	486	16	US-10-408-765A-1819
23	42.5	47.8	486	18	US-10-287-436A-595
24	42.5	47.8	778	11	US-09-833-245-912
25	42	47.2	122	16	US-10-425-115-198727
26	42	47.2	233	16	US-10-425-115-186741
27	42	47.2	432	15	US-10-424-599-192975
28	42	47.2	529	14	US-10-128-714-8425
29	42	47.2	626	14	US-10-081-872-132
30	42	47.2	626	15	US-10-385-305-132
31	42	47.2	820	16	US-10-684-424-44
32	42	47.2	1310	15	US-10-334-143-22
33	42	47.2	1999	10	US-09-978-244A-8
34	42	47.2	2090	9	US-09-736-969A-93
35	42	47.2	2090	9	US-09-737-246-2
36	42	47.2	2090	9	US-09-737-246-101
37	42	47.2	2090	9	US-09-736-960-90
38	42	47.2	2090	9	US-09-736-968A-107
39	42	47.2	2090	10	US-09-978-244A-28
40	41	46.1	64	16	US-10-425-115-303295
41	41	46.1	82	16	US-10-437-963-133052
42	41	46.1	86	16	US-10-425-115-191201
43	41	46.1	112	16	US-10-739-930-6954
44	41	46.1	133	16	US-10-408-765A-918
45	41	46.1	133	16	US-10-437-963-171387

#### ALIGNMENTS

#### RESULT 1

US-10-935-155A-4  
; Sequence 4, Application US/10935155A  
; Publication No. US20050037014A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-212  
; CURRENT APPLICATION NUMBER: US/10/935,155A  
; PRIOR FILING DATE: 2004-09-08  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 10/705,819  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: PCT/AU00/000070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P8  
US-10-935-155A-4

```
Query Match      100.0%; Score 89; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRSDRPVSYTMNRTS 17
      |||||
Db      1 PRSDRPVSYTMNRTS 17
      |||||

RESULT 2
US-10-705-819B-4
; Sequence 4, Application US/10705819B
; Publication No. US20050186220A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-201
; CURRENT APPLICATION NUMBER: US/10/705,819B
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: US 09/890650
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: PCT/AU00/00070
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: AU PP8533
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(17)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v
; OTHER INFORMATION: peptide p8
US-10-705-819B-4

Query Match      100.0%; Score 89; DB 18; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRSDRPVSYTMNRTS 17
      |||||
Db      1 PRSDRPVSYTMNRTS 17
      |||||

RESULT 3
US-09-951-061A-141
; Sequence 141, Application US/09951061A
; Publication No. US2003008220A1
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Tartaglia, James
; APPLICANT: Taylor, Jill
; APPLICANT: Gettig, Russell
; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMP VIRUS (CDV)
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE
; TITLE OF INVENTION: RECOMBINANTS
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell, Boehnen, Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/951,061A
; FILING DATE: 13-SEP-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/354,138
; FILING DATE: 15-JUL-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/224,657
; FILING DATE: 16-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,962
; FILING DATE: 08-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,867
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,283
; FILING DATE: 31-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/713,967
; FILING DATE: 11-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07,666,056
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-951-061A-141

Query Match      100.0%; Score 89; DB 10; Length 662;
Best Local Similarity 100.0%; Pred. No. 8e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PRSDRPVSYTMNRTS 17
      |||||
Db      50 PRSDRPVSYTMNRTS 66
      |||||

RESULT 4
US-10-670-695-36
; Sequence 36, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
```

; APPLICANT: Rosen, David K.  
; APPLICANT: Andrews, Janet S.  
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
; FILE REFERENCE: DI-9-1  
; CURRENT APPLICATION NUMBER: US/10/670,695  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 09/521,738  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 36  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: canine distemper virus  
US-10-670-695-36

Query Match 100.0%; Score 89; DB 15; Length 662;  
Best Local Similarity 100.0%; Pred. No. 8e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRSDRPVSYTMNRTS 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 50 PRSDRPVSYTMNRTS 66

## RESULT 5

US-10-935-155A-27  
; Sequence 27, Application US/10935155A  
; Publication No. US2005037014A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-212  
; CURRENT APPLICATION NUMBER: US/10/935,155A  
; CURRENT FILING DATE: 2004-09-08  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PF8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(662)  
; OTHER INFORMATION: Description of Artificial Sequence: full length canine distemper  
; OTHER INFORMATION: fusion protein  
US-10-935-155A-27

Query Match 100.0%; Score 89; DB 17; Length 662;  
Best Local Similarity 100.0%; Pred. No. 8e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRSDRPVSYTMNRTS 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 50 PRSDRPVSYTMNRTS 66

## RESULT 6

US-10-705-819B-27  
; Sequence 27, Application US/10705819B

; Publication No. US20050186220A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-201  
; CURRENT APPLICATION NUMBER: US/10/705,819B  
; CURRENT FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PF8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 662  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(662)  
; OTHER INFORMATION: Description of Artificial Sequence: full length canine distemper  
; OTHER INFORMATION: fusion protein  
US-10-705-819B-27

Query Match 100.0%; Score 89; DB 18; Length 662;  
Best Local Similarity 100.0%; Pred. No. 8e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PRSDRPVSYTMNRTS 17  
| | | | | | | | | | | | | | | | | | | | | |  
Db 50 PRSDRPVSYTMNRTS 66

## RESULT 7

US-10-437-963-104337  
; Sequence 104337, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 104337  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(112)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_101681C.1.pep  
US-10-437-963-104337

Query Match 52.8%; Score 47; DB 16; Length 112;  
Best Local Similarity 50.0%; Pred. No. 8.2;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 PRTSDRPVSYTMNTR 16  
| : : : : :  
Db 65 PSTSDKSVGYTTNKAK 80

## RESULT 8

US-10-437-963-134736  
; Sequence 134736, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 134736  
; LENGTH: 307  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(307)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_36480C.1.pep  
US-10-437-963-134736

Query Match 52.8%; Score 47; DB 16; Length 307;  
Best Local Similarity 52.9%; Pred. No. 25;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 PRTSDRPVSYTMNTRS 17  
| : : : : :  
Db 82 PPTSDKSVGYTTNKGTS 98

## RESULT 9

US-10-156-761-11575  
; Sequence 11575, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 11575  
; LENGTH: 297  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-11575

Query Match 51.7%; Score 46; DB 14; Length 297;  
Best Local Similarity 66.7%; Pred. No. 35;  
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 RTSDRPVSYTMNTR 16  
| : : : : :  
Db 11 RSSDRPHSETARRTR 25

## RESULT 10

US-10-437-963-168921  
; Sequence 168921, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 168921  
; LENGTH: 94  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(94)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_6738C.1.pep  
US-10-437-963-168921

Query Match 50.6%; Score 45; DB 16; Length 94;  
Best Local Similarity 57.1%; Pred. No. 14;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PRTSDRPVSYTMNR 14  
| : : : : :  
Db 64 PSTSDKSVGYTTNK 77

## RESULT 11

US-10-128-714-3066  
; Sequence 3066, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wendi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Ershkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3066  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-3066

Query Match 50.6%; Score 45; DB 14; Length 402;  
Best Local Similarity 52.9%; Pred. No. 70;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17  
Db 367 PRTTPTSLRRSRARS 383

## RESULT 12

US-10-425-115-341136  
; Sequence 341136, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 341136  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(79)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-425-115-341136

Query Match 49.4%; Score 44; DB 16; Length 49;  
Best Local Similarity 58.8%; Pred. No. 10;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNRTS 17  
Db 11 PRDSRRPSSYPFLATES 27

## RESULT 13

US-10-437-963-190524  
; Sequence 190524, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 190524  
; LENGTH: 50  
; TYPE: PRT

; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_86930C.1.pep  
US-10-437-963-190524

Query Match 49.4%; Score 44; DB 16; Length 50;  
Best Local Similarity 57.1%; Pred. No. 11;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNR 14  
Db 30 PPTSDKTGVYTNK 43

## RESULT 14

US-10-437-963-126154  
; Sequence 126154, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 126154  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(79)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_28729C.1.pep  
US-10-437-963-126154

Query Match 49.4%; Score 44; DB 16; Length 79;  
Best Local Similarity 57.1%; Pred. No. 17;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PRTSDRPVSYTMNR 14  
Db 64 PPTSDKSGVYTNK 77

## RESULT 15

US-10-437-963-108772  
; Sequence 108772, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966

```

; SEQ ID NO 108772
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12995C.1.pep
US-10-437-963-108772

Query Match      49.4%; Score 44; DB 16; Length 102;
Best Local Similarity 57.1%; Pred.No. 23;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      1 PRTSDRPVSYTMNR 14
      |||||
Db      64 PPTSDKSVGYTNK 77

Search completed: September 28, 2005, 18:26:01
Job time : 88 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 17:56:03 ; Search time 87 Seconds  
(without alignments)  
79.534 Million cell updates/sec

Title: US-10-705-819B-25

Perfect score: 88

Sequence: 1 SPKLLTFIASDTCPLV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1826554 seqs, 407025358 residues

Total number of hits satisfying chosen parameters: 1826554

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	100.0	17	17	US-10-935-155A-25
2	88	100.0	17	18	US-10-705-819B-25
3	88	100.0	438	15	US-10-267-682-93
4	88	100.0	438	15	US-10-267-748-93
5	88	100.0	662	10	US-09-951-061A-141
6	88	100.0	662	15	US-10-670-695-36
7	88	100.0	662	17	US-10-935-155A-27
8	88	100.0	662	18	US-10-705-819B-27
9	66	75.0	438	15	US-10-267-682-105
10	66	75.0	438	15	US-10-267-748-105
11	66	75.0	550	9	US-09-873-233A-25
12	65	73.9	49	17	US-10-935-155A-18
13	49	55.7	17	18	US-10-705-819B-24
14	49	55.7	17	18	US-10-705-819B-24
15	47.5	54.0	755	9	US-09-919-497-57
16	47.5	54.0	1780	15	US-10-115-479-10
17	47	53.4	372	15	US-10-369-493-7139
18	47	53.4	372	15	US-10-369-493-7359
19	47	53.4	372	15	US-10-369-493-7359
20	47	53.4	372	15	US-10-369-493-7359
21	44	50.0	418	16	US-10-425-115-295958
22	43.5	49.4	3063	14	US-09-918-715-257
23	43.5	49.4	3063	14	US-10-177-293-61
24	43.5	49.4	3063	14	US-10-177-293-63
25	43.5	49.4	3063	16	US-10-301-822-26
26	43.5	49.4	3063	16	US-10-474-794-257
27	43.5	49.4	3063	18	US-10-979-159-257
28	43.5	49.4	3118	18	US-10-220-335-287
29	43	48.9	59	18	US-10-450-763-59444
30	43	48.9	65	16	US-10-425-115-34172
31	43	48.9	91	16	US-10-425-115-322755
32	43	48.9	662	18	US-10-450-763-48063
33	43	48.9	1246	9	US-09-741-669-349
34	43	48.9	1246	15	US-10-369-493-831
35	42	47.7	76	17	US-10-743-280A-11
36	42	47.7	375	15	US-10-369-493-4948
37	42	47.7	376	15	US-10-369-493-7706
38	42	47.7	415	16	US-10-437-963-197942
39	42	47.7	481	16	US-10-767-701-41646
40	42	47.7	511	15	US-10-424-599-256019
41	42	47.7	552	15	US-10-369-493-12621
42	41.5	47.2	306	16	US-10-789-247-10
43	41	46.6	270	16	US-10-425-115-304421
44	41	46.6	298	15	US-10-425-114-69788
45	41	46.6	400	18	US-10-450-763-33079

#### RESULT 1

US-10-935-155A-25  
Sequence 25, Application US/10935155A  
Publication No. US20050037014A1  
GENERAL INFORMATION:  
APPLICANT: Jackson, David C  
APPLICANT: Ghosh, Souravi  
APPLICANT: Walker, John  
TITLE OF INVENTION: T Helper Cell Epitopes  
FILE REFERENCE: 47-212  
CURRENT APPLICATION NUMBER: US/10/935,155A  
CURRENT FILING DATE: 2004-09-08  
PRIOR APPLICATION NUMBER: US 09/890650  
PRIOR FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: US 10/705,819  
PRIOR FILING DATE: 2003-11-13  
PRIOR APPLICATION NUMBER: PCT/AU00/000070  
PRIOR FILING DATE: 2000-02-07  
PRIOR APPLICATION NUMBER: AU PP8533  
PRIOR FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: AU PQ2013  
PRIOR FILING DATE: 1999-08-04  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 25  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(17)  
OTHER INFORMATION: Description of Artificial Sequence:  
OTHER INFORMATION: peptide P75  
US-10-935-155A-25

synthetic canine distemper v

Query Match 100.0%; Score 88; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 9.5e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17  
|||||  
Db 1 SPDKLLTFIASDTCPLV 17

## RESULT 2

US-10-705-819B-25  
; Sequence 25, Application US/10705819B  
; Publication No. US20050186220A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-201  
; CURRENT APPLICATION NUMBER: US/10/705,819B  
; CURRENT FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU P8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU P02013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P75  
US-10-705-819B-25

Query Match 100.0%; Score 88; DB 18; Length 17;  
Best Local Similarity 100.0%; Pred. No. 9.5e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17  
|||||  
Db 1 SPDKLLTFIASDTCPLV 17

## RESULT 3

US-10-267-682-93  
; Sequence 93, Application US/10267682  
; Publication No. US2004003235A1  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; Wild, Carl T.  
; Barney, Shawn O.  
; Lambert, Dennis M.  
; Petteway, Stephen R.  
; Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TRANSMISSION  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York

COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/10/267,682  
APPLICATION NUMBER: US/10/267,682  
FILING DATE: 08-Oct-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,223A  
FILING DATE: 07-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 438 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-10-267-682-93

Query Match 100.0%; Score 88; DB 15; Length 438;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SPDKLLTFIASDTCPLV 17  
|||||  
Db 295 SPDKLLTFIASDTCPLV 311

## RESULT 4

US-10-267-748-93  
; Sequence 93, Application US/10267748  
; Publication No. US20040052820A1  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; Wild, Carl T.  
; Barney, Shawn O.  
; Lambert, Dennis M.  
; Petteway, Stephen R.  
; Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TRANSMISSION  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/10/267,748  
APPLICATION NUMBER: US/10/267,748  
FILING DATE: 08-Oct-2002  
CLASSIFICATION: <Unknown>



```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 93:
US-10-267-748-93

Query Match 100.0%; Score 88; DB 15; Length 438;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFFIASDTCPLV 17
Db 295 SPDKLLTFFIASDTCPLV 311

RESULT 5
US-09-951-061A-141
; Sequence 141, Application US/09951061A
; Publication No. US20030082204A1
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; APPLICANT: Tartaglia, James
; APPLICANT: Taylor, Jill
; APPLICANT: Gettig, Russell
; TITLE OF INVENTION: POXVIRUS - CANINE DISTEMPER VIRUS (CDV)
; TITLE OF INVENTION: RECOMBINANTS AND COMPOSITIONS AND METHODS EMPLOYING THE
; TITLE OF INVENTION: RECOMBINANTS
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell, Boehnen, Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/951.061A
; FILING DATE: 13-SEP-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/354,138
; FILING DATE: 15-JUL-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/224,657
; FILING DATE: 16-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,962
; FILING DATE: 08-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/776,867
; FILING DATE: 23-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/938,283
; FILING DATE: 31-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/621,614
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/105,483
; FILING DATE: 12-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,951
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/713,967
; FILING DATE: 11-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07,666,056
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2860
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-951-061A-141

Query Match 100.0%; Score 88; DB 10; Length 662;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFFIASDTCPLV 17
Db 519 SPDKLLTFFIASDTCPLV 535

RESULT 6
US-10-670-695-36
; Sequence 36, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 662
; TYPE: PRT
; ORGANISM: canine distemper virus
; US-10-670-695-36

Query Match 100.0%; Score 88; DB 15; Length 662;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFFIASDTCPLV 17
```

```

Db          519 SPDKLLTFIASDTCPLV 535
|||||
RESULT 7
US-10-935-155A-27
; Sequence 27, Application US/10935155A
; Publication No. US20050037014A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-212
; CURRENT APPLICATION NUMBER: US/10/935,155A
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: US 09/890650
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 10/705,819
; PRIOR FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00070
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: AU PP8533
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(662)
; OTHER INFORMATION: Description of Artificial Sequence: full length canine distemper
; OTHER INFORMATION: fusion protein
US-10-935-155A-27

Query Match          100.0%; Score 88; DB 17; Length 662;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 SPDKLLTFIASDTCPLV 17
|||||
Db          519 SPDKLLTFIASDTCPLV 535
|||||
RESULT 8
US-10-705-819B-27
; Sequence 27, Application US/10705819B
; Publication No. US20050186220A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-201
; CURRENT APPLICATION NUMBER: US/10/705,819B
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00070
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 09/890650
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: AU PP8533
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: AU PQ2013
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 662
; TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(662)
; OTHER INFORMATION: Description of Artificial Sequence: full length canine distemper
; OTHER INFORMATION: fusion protein
US-10-705-819B-27

Query Match          100.0%; Score 88; DB 18; Length 662;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1 SPDKLLTFIASDTCPLV 17
|||||
Db          519 SPDKLLTFIASDTCPLV 535
|||||
RESULT 9
US-10-267-682-105
; Sequence 105, Application US/10267682
; Publication No. US20040033235A1
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/267,682
; FILING DATE: 08-Oct-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-10-267-682-105

Query Match          75.0%; Score 66; DB 15; Length 438;
Best Local Similarity 68.8%; Pred. No. 0.018;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```
Qy 2 PDKLLTFIASDTCPLV 17
Db 296 PDKILTYIAADHCPV 311

RESULT 10
US-10-267-748-105
; Sequence 105, Application US/10267748
; Publication No. US20040052820A1
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; ; Matthews, Thomas J.
; ; Wild, Carl T.
; ; Barney, Shawn O.
; ; Lambert, Dennis M.
; ; Petteway, Stephen R.
; ; Langlois, Alphonse J.
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/267,748
; FILING DATE: 08-Oct-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 438 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 105:
US-10-267-748-105

Query Match 75.0%; Score 66; DB 15; Length 438;
Best Local Similarity 68.8%; Pred. No. 0.018;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDKLLTFIASDTCPLV 17
Db 296 PDKILTYIAADHCPV 311

RESULT 11
US-09-873-233A-20
; Sequence 20, Application US/09873233A
; Patent No. US20020146434A1
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; TITLE OF INVENTION: GENE CODING FOR THE MEASLES VIRUS MUTANT ANTIGEN
; FILE REFERENCE: 0216-0451P
; CURRENT APPLICATION NUMBER: US/09/873,233A
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(550)
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-873-233A-20

Query Match 75.0%; Score 66; DB 9; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.023;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDKLLTFIASDTCPLV 17
Db 408 PDKILTYIAADHCPV 423

RESULT 12
US-09-873-233A-18
; Sequence 18, Application US/09873233A
; Patent No. US20020146434A1
; GENERAL INFORMATION:
; APPLICANT: UEDA, Shigeharu
; APPLICANT: WATANABE, Michiko
; APPLICANT: KAWANISHI, Hitomi
; TITLE OF INVENTION: GENE CODING FOR THE MEASLES VIRUS MUTANT ANTIGEN
; FILE REFERENCE: 0216-0451P
; CURRENT APPLICATION NUMBER: US/09/873,233A
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 18
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Measles virus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(550)
; OTHER INFORMATION: any n or Xaa = Unknown
US-09-873-233A-18

Query Match 73.9%; Score 65; DB 9; Length 550;
Best Local Similarity 68.8%; Pred. No. 0.034;
Matches 11; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PDKLLTFIASDTCPLV 17
Db 408 PDKILTYIAADHCPV 423

RESULT 13
US-10-935-155A-24
; Sequence 24, Application US/10935155A
; Publication No. US20050037014A1
; GENERAL INFORMATION:
; APPLICANT: Jackson, David C
; APPLICANT: Ghosh, Souravi
; APPLICANT: Walker, John
; TITLE OF INVENTION: T Helper Cell Epitopes
; FILE REFERENCE: 47-212
; CURRENT APPLICATION NUMBER: US/10/935,155A
; CURRENT FILING DATE: 2004-09-08
```

; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: US 10/705,819  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P74  
US-10-935-155A-24

Query Match 55.7%; Score 49; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIA 10  
| | | | | | | |  
Db 8 SPDKLLTFIA 17

RESULT 14  
US-10-705-819B-24  
; Sequence 24, Application US/10705819B  
; Publication No. US20050186220A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackson, David C  
; APPLICANT: Ghosh, Souravi  
; APPLICANT: Walker, John  
; TITLE OF INVENTION: T Helper Cell Epitopes  
; FILE REFERENCE: 47-201  
; CURRENT APPLICATION NUMBER: US/10/705,819B  
; PRIOR FILING DATE: 2003-11-13  
; PRIOR APPLICATION NUMBER: US 09/890650  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: PCT/AU00/00070  
; PRIOR FILING DATE: 2000-02-07  
; PRIOR APPLICATION NUMBER: AU PP8533  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: AU PQ2013  
; PRIOR FILING DATE: 1999-08-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(17)  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic canine distemper v  
; OTHER INFORMATION: peptide P74  
US-10-705-819B-24

Query Match 55.7%; Score 49; DB 18; Length 17;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SPDKLLTFIA 10  
| | | | | | | |  
Db 8 SPDKLLTFIA 17

RESULT 15  
US-09-919-497-57  
; Sequence 57, Application US/09919497  
; Patent No. US20020106662A1  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 57  
; LENGTH: 755  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-57

Query Match 54.0%; Score 47.5; DB 9; Length 755;  
Best Local Similarity 55.6%; Pred. No. 43;  
Matches 10; Conservative 4; Mismatches 1; Indels 3; Gaps 1;

Qy 3 DKLLTFI---ASDTCPLV 17  
| : | | : | | | | : |  
Db 178 DELITFVCETASATCPVV 195

Search completed: September 28, 2005, 18:26:04  
Job time : 88 secs